Non-Linear Regression with Multidimensional Indices

Naveen K. Bansal, G. G. Hamedani

Department of Mathematics, Statistics, and Computer Science

Marquette University

Milwaukee, WI 53233

and

Hao Zhang

Program in Statistics

Washington State University

Pullman, WA 99164

Summary

We consider non-linear regression model when the index variable is multidimensional. Sufficient conditions on the non-linear function are given under which the least squares estimators are strongly consistent and asymptotically normally distributed. These sufficient conditions are satisfied by harmonic type functions, which are also of interest in one dimensional index case where Wu's (1981) and Jennrich's (1969) sufficient conditions are not applicable.

1 Introduction.

Statistical modeling with multidimensional indices is an important problem in spatial or tempro-spatial process, in signal processing, and in texture modeling. For example, suppose the problem is to model the growth of vegetation in a particular farm over a period of time. This requires the modeling with three dimensional indices. For examples in signal processing, see Rao, Zhao and Zhou (1994) and McClellan (1982). For examples in texture modeling, see Francos, Meiri and Porot (1993), Yuan and Subba Rao (1993), and Mandrekar and Zhang (1996). Most of the models in these areas are non-linear regression models. There has been extensive work in the literature on non-linear regression models with one dimensional index (Wu (1981), Jennrich (1969), and Gallant (1987)). A non-linear regression model with one dimensional index can be defined as follows

$$y_t = f(x_t, \boldsymbol{\theta}) + \epsilon_t, \quad t = 1, 2, \ldots, n, \tag{1}$$

where the observed data is $\{y_t,\ t = 1 \ldots n\}$, $x_t,\ t = 1 \ldots n$ are some known constants, $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subseteq \mathbb{R}^p$ is an unknown parameter vector, $\epsilon_t,\ t = 1 \ldots n$ are the random errors, and f is a known function. The problem of estimating $\boldsymbol{\theta}$ has been investigated quite extensively in literature. Wu (1981) and Jennrich (1969) gave some sufficient conditions based on the function f and the design to establish certain asymptotic properties of the least squares estimators. These conditions, however, are not satisfied if the function f is of harmonic type (Kundu (1993)).

In the present work, we consider an extension to (1) with multidimensional indices,

$$y_{\mathbf{t}} = f(\mathbf{x_t}, \boldsymbol{\theta}) + \epsilon_{\mathbf{t}}, \quad \mathbf{t} \leq \mathbf{n}, \tag{2}$$

where $\mathbf{t} = (t_1, t_2, \ldots, t_k)$ and $\mathbf{n} = (n_1, n_2, \ldots, n_k) \in \mathbb{N}^k$ (set of k dimensional non-negative integer values), $\leq$ denotes the partial ordering, i.e., for $\mathbf{m} = (m_1, m_2, \ldots, m_k) \in \mathbb{N}^k$ and $\mathbf{n} = (n_1, n_2, \ldots, n_k) \in \mathbb{N}^k, \mathbf{m} \leq \mathbf{n}$ if $m_i \leq n_i$ for $i = 1, \ldots, k$, $\left\{\epsilon_{\mathbf{t}},\ \mathbf{t} \in \mathbb{N}^k\right\}$ is an independent field of random variables such that

$$E(\epsilon_{\mathbf{t}}) = 0 \quad \text{and} \quad \mathrm{Var}\,(\epsilon_{\mathbf{t}}) = \sigma^2 \quad \forall\, \mathbf{t} \in \mathbb{N}^k, \tag{3}$$

$\boldsymbol{\theta} \in \boldsymbol{\Theta} \subseteq \mathbb{R}^p$ is a parameter vector, $\{\mathbf{x_t},\ \mathbf{t} \in \mathbb{N}^k\}$ a set of known constant vectors, and f is a known non-linear function.

Remark. For signal processing models (see Rao, Zhao and Zhou (1994) for an example), $y_{\mathbf{t}}$, $f(\cdot, \cdot)$ and $\epsilon_{\mathbf{t}}$ are complex valued. Here for notational convenience we only assume them to be real valued.

A natural choice of estimating $\boldsymbol{\theta}$ is by least squares methods, i.e., by minimizing

$$Q_{\mathbf{n}}(\boldsymbol{\theta}) = \sum_{\mathbf{t} \leq \mathbf{n}} |y_{\mathbf{t}} - f(\mathbf{x_t},\ \boldsymbol{\theta})|^2. \tag{4}$$

Here we will not deal with the numerical method of estimating $\boldsymbol{\theta}$. Our aim will be to find sufficient conditions on function $f(\cdot, \cdot)$ so that the least squares estimators are strongly consistent and are asymptotically normally distributed as $|\mathbf{n}| = \prod_{i=1}^{k} n_i \to \infty$. Note that consistency and asymptotic normality as $|\mathbf{n}| \to \infty$ provide much stronger results than consistency and asymptotic normality as $\min(n_1, n_2, \ldots, n_k) \to \infty$, as assumed, for example, in Rao, Zhao and Zhou. Our results will also be of interest in one dimensional case since they can be applied to harmonic type functions which do not satisfy Wu's and Jennrich's sufficient conditions (Kundu (1993)).

To illustrate, we will consider the following example which can be used to model textures (Yuan and Subba Rao (1993) and Mandrekar and Zhang (1996)):

$$y_{\mathbf{t}} = \sum_{k=1}^{p} \alpha_k \cos(t_1 \lambda_{1k} + t_2 \lambda_{2k}) + \epsilon_{\mathbf{t}}, \tag{5}$$

where $\mathbf{t} = (t_1, t_2)$, α_k's are real unknown parameters, and $\lambda_{1k}, \lambda_{2k}$ are unknown parameters in $[0, \pi]$.

This model will be taken up in details in Section 2 as well as in Section 3. In Section 2, we will present sufficient conditions in terms of function f to establish strong consistency of $\widehat{\boldsymbol{\theta}}_{\mathbf{n}}$. In Section 3, sufficient conditions for the asymptotic normality of $\widehat{\boldsymbol{\theta}}_{\mathbf{n}}$ will be given and we will obtain its asymptotic distribution.

2 Strong Consistency.

Let $\boldsymbol{\theta}_0$ be the true parameter vector. Our objective is to prove the strong consistency of $\widehat{\boldsymbol{\theta}}_{\mathbf{n}}$ in the sense that $\widehat{\boldsymbol{\theta}}_{\mathbf{n}} \longrightarrow \boldsymbol{\theta}_0$ a.s. as $|\mathbf{n}| \to \infty$. Note that if $\{y_{\mathbf{t}},\ \mathbf{1} \le \mathbf{t} \le \mathbf{n}\}$, where

$\mathbf{1} = (1, 1, \ldots, 1) \in \mathbb{N}^k$, is the observed data, then the total number of observations is $|\mathbf{n}|$. To prove the strong consistency, we need the following lemma which is similar to lemma 1 of Wu (1981).

Lemma 2.1. *Let $\{Q_{\mathbf{n}}(\boldsymbol{\theta})\}$ be a field of measurable functions such that*

$$\liminf_{|\mathbf{n}|\to\infty} \inf_{|\boldsymbol{\theta}-\boldsymbol{\theta}_0|\geq\delta} \frac{1}{|\mathbf{n}|} \left(Q_{\mathbf{n}}(\boldsymbol{\theta}) - Q_{\mathbf{n}}(\boldsymbol{\theta}_0)\right) > 0 \quad a.s. \tag{6}$$

for every $\delta > 0$. Then $\widehat{\boldsymbol{\theta}}_{\mathbf{n}}$ which mimnimizes $Q_{\mathbf{n}}(\boldsymbol{\theta})$ converges to $\boldsymbol{\theta}_0$ a.s. as $|\mathbf{n}| \to \infty$.

Proof: Suppose $\widehat{\boldsymbol{\theta}}_{\mathbf{n}} \not\to \boldsymbol{\theta}_0$ a.s. as $|\mathbf{n}| \to \infty$. Then there exists a subsequence $\{\mathbf{n}_s,\ s \geq 1\}$ and $\delta > 0$ such that $\left|\widehat{\boldsymbol{\theta}}_{\mathbf{n}_s} - \boldsymbol{\theta}_0\right| \geq \delta$ for all $s \geq 1$ with positive probability. Thus, from (6),

$$Q_{\mathbf{n}_s}\left(\widehat{\boldsymbol{\theta}}_{\mathbf{n}_s}\right) - Q_{\mathbf{n}_s}(\boldsymbol{\theta}_0) > 0$$

with positive probability for all $s \geq M$, for some $M > 0$. This contradicts the fact that $\widehat{\boldsymbol{\theta}}_{\mathbf{n}_s}$ is a least squares estimator.

We now make the following assumptions:

Assumption 1: The parameter space $\boldsymbol{\Theta}$ is compact.

Assumption 2: The function $f(\mathbf{x_t}, \cdot)$ satisfies

(*i*) $|f(\mathbf{x_t}, \boldsymbol{\theta}_1) - f(\mathbf{x_t}, \boldsymbol{\theta}_2)| \leq a_{\mathbf{t}}|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2|$ for all $\boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_2$,

where $a_{\mathbf{t}} > 0$, $\mathbf{t} \in \mathbb{N}^k$ are some constants such that $\sum_{\mathbf{t}\leq\mathbf{n}} a_{\mathbf{t}} = o(|\mathbf{n}|^3)$, (*ii*) $\sup_{\mathbf{t}\in\mathbb{N}^k,\boldsymbol{\theta}\in\boldsymbol{\Theta}} |f(\mathbf{x_t},\boldsymbol{\theta})| \le M_0$ for some $M_0 > 0$.

Assumption 3:

$$\lim \inf_{|\mathbf{n}|\to\infty} \inf_{|\boldsymbol{\theta}-\boldsymbol{\theta}_0|\ge\delta} \frac{1}{|\mathbf{n}|} \sum_{\mathbf{t}\le\mathbf{n}} [f(\mathbf{x_t},\boldsymbol{\theta}) - f(\mathbf{x_t},\boldsymbol{\theta}_0)]^2 > 0.$$

Let $\widehat{\boldsymbol{\theta}}_{\mathbf{n}}$ denote the least squares estimator which minimizes (4).

Theorem 2.2. *Under Assumptions 1 – 3, and (3),* $\widehat{\boldsymbol{\theta}}_{\mathbf{n}} \to \boldsymbol{\theta}_0$ *a.s. as* $|\mathbf{n}| \to \infty$.

Proof: Observe that

$$\frac{1}{|\mathbf{n}|}\left(Q_{\mathbf{n}}(\boldsymbol{\theta}) - Q_{\mathbf{n}}(\boldsymbol{\theta}_0)\right) = \frac{1}{|\mathbf{n}|}\sum_{\mathbf{t}\le\mathbf{n}} [f(\mathbf{x_t},\boldsymbol{\theta}) - f(\mathbf{x_t},\boldsymbol{\theta}_0)]^2 + \frac{1}{|\mathbf{n}|}\sum_{\mathbf{t}\le\mathbf{n}} [f(\mathbf{x_t},\boldsymbol{\theta}) - f(\mathbf{x_t},\boldsymbol{\theta}_0)]\,\epsilon_{\mathbf{t}}.$$

From Assumption 3, the first term on the RHS of the above equality is positive. Thus, from Lemma 2.1, the consistency of $\widehat{\boldsymbol{\theta}}_{\mathbf{n}}$ follows if we prove the following

$$\sup_{|\boldsymbol{\theta}-\boldsymbol{\theta}_0|\ge\delta} \frac{1}{|\mathbf{n}|}\left|\sum_{\mathbf{t}\le\mathbf{n}} [f(\mathbf{x_t},\boldsymbol{\theta}) - f(\mathbf{x_t},\boldsymbol{\theta}_0)]\,\epsilon_{\mathbf{t}}\right| \to 0 \quad a.s. \quad as \quad |\mathbf{n}| \to \infty. \tag{7}$$

Now, since in view of Assumption 2, the function $d(\mathbf{x_t},\boldsymbol{\theta}) = f(\mathbf{x_t},\boldsymbol{\theta}) - f(\mathbf{x_t},\boldsymbol{\theta}_0)$ is bounded, i.e.,

$$\sup_{\mathbf{t}\in\mathbb{N}^k,\, \boldsymbol{\theta}\in\boldsymbol{\Theta}} |d(\mathbf{x_t},\ \boldsymbol{\theta})| < \infty,$$

and since it satisfies the condition (**??**) bleow, (7) follows from the following Lemma.

Lemma 2.3. *Let* $g(\mathbf{x_t},\cdot)$ *be such that* *(i)* $|g(\mathbf{x_t},\, \boldsymbol{\theta}_1) - g(\mathbf{x_t},\, \boldsymbol{\theta}_2)| \leq a_{\mathbf{t}} |\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2|, \quad \textit{for all} \quad \boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_2$ (8)

where $\sum_{\mathbf{t} \leq \mathbf{n}} a_{\mathbf{t}} = o(|\mathbf{n}|^3)$,

(ii) $\sup_{\mathbf{t} \in \mathbb{N}^k, \boldsymbol{\theta} \in \boldsymbol{\Theta}} |g(\mathbf{x_t}, \boldsymbol{\theta})| = B < \infty,$ (9)

and let $\{\epsilon_{\mathbf{t}},\, \mathbf{t} \in \mathbb{N}^k\}$ *be i.i.d. real valued random field with mean 0 and variance* σ^2*, then*

$$\sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \left| \frac{1}{|\mathbf{n}|} \sum_{\mathbf{t} \leq \mathbf{n}} g(\mathbf{x_t}, \boldsymbol{\theta})\, \epsilon_{\mathbf{t}} \right| \longrightarrow 0 \quad a.s. \quad as \quad |\mathbf{n}| \to \infty.$$

Proof: To prove this, we use the following result of Mikosch and Norvaisa (1987) on Banach space valued random field.

Let $\left\{ X_{\mathbf{n}}, \mathbf{n} \in \mathbb{N}^k \right\}$ *be a field of Banach space valued random variables with a norm* $\| \cdot \|$. *Let* $\{a_{\mathbf{n}}, \mathbf{n} \in \mathbb{N}^k\}$ *be a field of positive numbers such that* $\lim_{|\mathbf{n}| \to \infty} a_{\mathbf{n}} = \infty$*, it satisfies the conditions* $A - D$ *of Mikosch and Norvaisa (1987), and*

$$\{a_{\mathbf{n}}\} \in F_2 = \left(\{a_{\mathbf{n}}\} : \sum_{j=q}^{\infty} j^{-3} |A_j| = O\left(q^{-2}\, |A_q|\right),\ q \to \infty \right), \tag{10}$$

where $A_j = \{\mathbf{n} : a_{\mathbf{n}} \leq j\}$*, and* $|A_j|$ *is the cardinality of* A_j.

If there exists a constant $c > 0$ *and a positive random variable* X *such that*

$$\sup_{\mathbf{n} \in \mathbb{N}^k} P\left(\| X_{\mathbf{n}} \| > x\right) \leq c\, P(X > x) \quad \forall\, x > 0, \tag{11}$$

and

$$\sum_{\mathbf{n} \in \mathbb{N}^k} P\left(X > a_{\mathbf{n}}\right) < \infty, \tag{12}$$

then $\lim_{|\mathbf{n}|\to\infty} \frac{1}{a_{\mathbf{n}}} \sum_{\mathbf{j}\le\mathbf{n}} X_{\mathbf{j}} = 0$ *a.s. is equivalent to*

$$\lim_{|\mathbf{n}|\to\infty} \frac{1}{a_{\mathbf{n}}} \sum_{\mathbf{j}\le\mathbf{n}} X_{\mathbf{j}} = 0 \quad \textit{in probability.} \tag{13}$$

We define $X_{\mathbf{t}} = g(\mathbf{x}_{\mathbf{t}}, \boldsymbol{\theta})\, \epsilon_{\mathbf{t}}$. Then $\{X_{\mathbf{t}},\ \mathbf{t} \in \mathbb{N}^k\}$ is a field of $C(\boldsymbol{\theta})$-valued random variables, where $C(\Theta)$ is the space of continuous function on the compact metric space with the sup norm. Define $a_{\mathbf{n}} = |\mathbf{n}|$. It is easy to see that $\{a_{\mathbf{n}},\ \mathbf{n} \in \mathbb{N}^k\}$ satisfy the conditions A – D of Mikasch and Norvaisa (1987).

Now, from the above mentioned result of Mikasch and Norvaisa, it suffices to show that $\{a_{\mathbf{n}},\ \mathbf{n} \in \mathbb{N}^k\}$ and $\{X_{\mathbf{n}},\ \mathbf{n} \in \mathbb{N}^k\}$ satisfy (10) – (13).

To show (10), we need to prove

$$\sum_{j=q}^{\infty} j^{-3}\, |A_j| = O(q^{-2}\, |A_q|), \quad \text{as } q \to \infty. \tag{14}$$

Using mathematical induction, it can be shown that

$$\int \cdots \int_{1\le y_1\cdot y_2 \ldots y_k \le j} dy_1 \ldots dy_k = j \left[\frac{(\ln j)^{k-1}}{(k-1)!} - \frac{(\ln j)^{k-2}}{(k-2)!} \right].$$

Therefore

$$|A_j| \sim j(\ln j)^{k-1} \quad \text{as} \quad j \to \infty. \tag{15}$$

Thus

$$\begin{aligned} \sum_{j=q}^{\infty} j^{-3} |A_j| \quad &\sim \quad \sum_{j=q}^{\infty} j^{-3} j\, (\ln j)^{k-1} \\ &\sim \quad \int_q^{\infty} y^{-2}\, (\ln y)^{k-1}\, dy. \end{aligned} \tag{16}$$

(The notation $b_j \sim c_j$ means $b_j/c_j = O(1),\ c_j/b_j = O(1)$, i.e., as $j \to \infty$ $M_1 c_j \le b_j \le M_2 c_j$ for some $M_1, M_2 > 0$.)

Now, since

$$0 < y^{-2} \left[(\ln y)^{k-1} - (\ln M)^{k-1} \right] < y^{-1.5} \quad \forall \quad y > M,$$

where M is sufficiently large, we have

$$0 < \int_M^\infty y^{-2} (\ln y)^{k-1}\, dy - (\ln M)^{k-1} \int_M^\infty y^{-2}\, dy < \infty.$$

We now conclude, from (15) and (16), that

$$\sum_{j=q}^{\infty} j^{-3} |A_j| \ \sim \ \int_q^\infty y^{-2} (\ln y)^{k-1}\, dy \sim q^{-1} (\ln q)^{k-1}$$

$$\sim \ O(q^{-2} |A_q|), \quad \text{as} \quad q \to \infty.$$

This proves (14).

To prove (11) and (12), note that, since $g(\mathbf{x_t},\ \boldsymbol{\theta})$ is bounded, for any $x > 0$,

$$P(\| X_{\mathbf{t}} \| > x) = P\left(\sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \left| g\left(\mathbf{x_t},\ \boldsymbol{\theta}\right) \right| |\epsilon_{\mathbf{t}}| > x \right) \le P(R|\epsilon_{\mathbf{1}}| > x), \tag{17}$$

where R is the upper bound of $|g(\cdot,\cdot)|$. Now, substituting $X = R\,|\epsilon_{\mathbf{1}}|$, we get

$$\sum_{\mathbf{n} \in \mathbb{N}^k} P(|X| > |\mathbf{n}|) \le \sum_{\mathbf{n} \in \mathbb{N}^k} \frac{R^2 \sigma^2}{|\mathbf{n}|^2} < \infty. \tag{18}$$

The inequalities (17) and (18) prove (11) and (12) respectively.

To prove (13), we first note that, from (8), $X_{\mathbf{t}}$, $\mathbf{t} \le \mathbf{n}$ are independent random variables with mean zero belonging to a subspace of $C(\mathbf{\Theta})$,

$$Lip(C(\mathbf{\Theta})) = \{h \in C(\mathbf{\Theta}) : |h(\boldsymbol{\theta}_1) - h(\boldsymbol{\theta}_2)| < A|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2)| \text{ for all } \boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_2\},$$

where A is some positive constant.

Therefore from the inequality (6) (of Appendix) of Wu (1981), we have

$$E||\sum_{\mathbf{t}\leq\mathbf{n}} X_{\mathbf{t}}||^2 \quad \leq \quad K\sum_{\mathbf{t}\leq\mathbf{n}} E||X_{\mathbf{t}}||_L^2, \quad \text{for some } K < \infty, \tag{19}$$

where, for $h \in Lip(C(\mathbf{\Theta}))$, the norm $||\cdot||_L$ is defined as

$$||h||_L = \sup_{\boldsymbol{\theta}_1\neq\boldsymbol{\theta}_2} \frac{|h(\boldsymbol{\theta}_1) - h(\boldsymbol{\theta}_2)|}{|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2)|} + |h(\boldsymbol{\theta}_a)|, \tag{20}$$

for $\boldsymbol{\theta}_a$ some fixed point in $\mathbf{\Theta}$.

Now, from (19), for any $\delta > 0$, we have

$$\begin{aligned} P\left(\frac{1}{|\mathbf{n}|}\left\|\sum_{\mathbf{t}\leq\mathbf{n}} X_{\mathbf{t}}\right\| > \delta\right) &\leq \frac{E\left(\left\|\sum_{\mathbf{t}\leq\mathbf{n}} X_{\mathbf{t}}\right\|^2\right)}{|\mathbf{n}|^2\,\delta^2} \\ &= \frac{K\sum_{\mathbf{t}\leq\mathbf{n}} E\,\|X_{\mathbf{t}}\|_L^2}{|\mathbf{n}|^2\,\delta^2} \\ &= \frac{K\sum_{\mathbf{t}\leq\mathbf{n}} \|\,g(\mathbf{x}_{\mathbf{t}},\cdot)\|_L^2\;\; E\left(\epsilon_{\mathbf{t}}^2\right)}{|\mathbf{n}|^2\,\delta^2} \\ &= \frac{K\sigma^2\sum_{\mathbf{t}\leq\mathbf{n}} \|\,g(\mathbf{x}_{\mathbf{t}},\cdot)\|_L^2}{|\mathbf{n}|^2\,\delta^2}. \end{aligned} \tag{21}$$

From (8), (9) and (20), we have

$$||g(\mathbf{x}_{\mathbf{t}},\cdot)||_L \quad \leq \quad A_{\mathbf{t}} + B \quad \text{for all} \quad \mathbf{t} \in \mathbb{N}^k.$$

Thus, from (27),

$$P\left(\frac{1}{|\mathbf{n}|}\left\|\sum_{\mathbf{t}\le\mathbf{n}} X_{\mathbf{t}}\right\| > \delta\right) \le \frac{2K\sigma^2\left(\sum_{\mathbf{t}\le\mathbf{n}} A_{\mathbf{t}}^2 + \|\mathbf{n}\|B\right)}{|\mathbf{n}|^2\,\delta^2} \to 0 \text{as } |\mathbf{n}| \to \infty.$$

This completes the proof of Lemma 2.3 and thus the proof of Theorem 2.2.

We now show that the least squares estimators for model (5) are strongly consistent under the assumption (3). For notational convenience, we assume that $p = 1$ and deal only with

$$y_{\mathbf{t}} = \alpha\cos(\lambda_1 t_1 + \lambda_2 t_2) + \epsilon_{\mathbf{t}}, \quad \mathbf{1} \le \mathbf{t} \le \mathbf{n}, \tag{22}$$

where $\lambda_1, \lambda_2 \in [0, \pi]$. Further, we assume that $|\alpha| \le M < \infty$, for some $M > 0$. This is a reasonable assumption since α represents the amplitute of the waves.

To prove the strong consistency, via Theorem 2.2, it suffices to show that Assumptions 1 – 3 are satisfied for this model.

We let $\boldsymbol{\theta} = (\alpha, \lambda_1, \lambda_2)^T$and $\boldsymbol{\theta}_0 = (\alpha_o, \lambda_{1o}, \lambda_{2o})^T \in \boldsymbol{\Theta} = [-M, M] \times [0, \pi]^2$. Clearly Assumptions 1 and 2 are satisfied by taking $A_{\mathbf{t}} = 1$ and $M_o = M$. To varify Assumption 3, note that

$$\frac{1}{|\mathbf{n}|}\sum_{\mathbf{t}\le\mathbf{n}}\left[\alpha\, cos(\lambda_1 t_1 + \lambda_2 t_2) - \alpha_o\, cos(\lambda_{10} t_1 + \lambda_{20} t_2)\right]^2$$

$$= \frac{1}{|\mathbf{n}|}\sum_{\mathbf{t}\le\mathbf{n}}\left[\alpha\frac{e^{i(\lambda_1 t_1 + \lambda_2 t_2)} + e^{-i(\lambda_1 t_1 + \lambda_2 t_2)}}{2} - \alpha_o\frac{e^{i(\lambda_{1o} t_1 + \lambda_{2o} t_2)} + e^{-i(\lambda_{1o} t_1 + \lambda_{2o} t_2)}}{2}\right]^2$$

$$= \frac{\alpha^2}{4}\left[2 + \frac{1}{|\mathbf{n}|}\sum_{\mathbf{t}\le\mathbf{n}} e^{2i(\lambda_1 t_1 + \lambda_2 t_2)} + \frac{1}{|\mathbf{n}|}\sum_{\mathbf{t}\le\mathbf{n}} e^{-2i(\lambda_1 t_1 + \lambda_2 t_2)}\right]$$

$$+\frac{\alpha_o^2}{4}\left[2+\frac{1}{|\mathbf{n}|}\sum_{\mathbf{t}\le\mathbf{n}} e^{2i(\lambda_{1o}t_1+\lambda_{2o}t_2)}+\frac{1}{|\mathbf{n}|}\sum_{\mathbf{t}\le\mathbf{n}} e^{-2i(\lambda_{1o}t_1+\lambda_{2o}t_2)}\right]$$

$$-\frac{\alpha\alpha_o}{2|\mathbf{n}|}\sum_{\mathbf{t}\le\mathbf{n}}[e^{i((\lambda_1+\lambda_{1o})t_1+(\lambda_2+\lambda_{2o})t_2)}+e^{-i((\lambda_1+\lambda_{1o})t_1+(\lambda_2+\lambda_{2o})t_2)}$$

$$+e^{i((\lambda_1-\lambda_{1o})t_1+(\lambda_2-\lambda_{2o})t_2)}+e^{-i(\lambda_1-\lambda_{1o})t_1+(\lambda_2-\lambda_{2o})t_2}] \tag{23}$$

Now, since, for all ω_1 and $\omega_2 \in (0, 2\pi)$,

$$\sum_{\mathbf{t}\le\mathbf{n}} e^{i(\omega_1 t_1+\omega_2 t_2)} = e^{i(\omega_1+\omega_2)}\frac{(1-e^{in_1\omega_1})(1-e^{in_2\omega_2})}{(1-e^{i\omega_1})(1-e^{i\omega_2})},$$

the Assumption 3 follows from (23).

3 Asymptotic Normality.

Observe that the least squares estimator $\widehat{\boldsymbol{\theta}}_{\mathbf{n}}$ is a zero of $Q'_{\mathbf{n}}(\boldsymbol{\theta}) = 0$. (Primes are used to denote the derivatives with respect to $\boldsymbol{\theta}$). Thus expanding $Q'_{\mathbf{n}}(\boldsymbol{\theta})$ about $\boldsymbol{\theta}_0$ (true parameter value) and evaluating it at $\widehat{\boldsymbol{\theta}}_{\mathbf{n}}$, we get

$$0 = Q'_{\mathbf{n}}(\boldsymbol{\theta}_0) + Q''_{\mathbf{n}}(\widehat{\boldsymbol{\theta}}_{*\mathbf{n}})\left(\widehat{\boldsymbol{\theta}}_{\mathbf{n}} - \boldsymbol{\theta}_0\right), \tag{24}$$

where $\widehat{\boldsymbol{\theta}}_{*\mathbf{n}} = h\boldsymbol{\theta}_0 + (1-h)\,\widehat{\boldsymbol{\theta}}_{\mathbf{n}}$ for some $0 \le h \le 1$.

Note that

$$Q'_{\mathbf{n}}(\boldsymbol{\theta}) = -2\sum_{\mathbf{t}\le\mathbf{n}} (y_{\mathbf{t}} - f(\mathbf{x}_{\mathbf{t}},\boldsymbol{\theta}))\ f'(\mathbf{x}_{\mathbf{t}},\boldsymbol{\theta}) \tag{25}$$

$$Q''_{\mathbf{n}}(\boldsymbol{\theta}) = -2\sum_{\mathbf{t}\le\mathbf{n}} (y_{\mathbf{t}} - f(\mathbf{x}_{\mathbf{t}},\boldsymbol{\theta}))\ f''(\mathbf{x}_{\mathbf{t}},\boldsymbol{\theta}) + 2\sum_{\mathbf{t}\le\mathbf{n}} [f'(\mathbf{x}_{\mathbf{t}},\boldsymbol{\theta})]\,[f'(\mathbf{x}_{\mathbf{t}},\boldsymbol{\theta})]^T \tag{26}$$

$$
\begin{aligned}
= \quad & -2\sum_{\mathbf{t}\le\mathbf{n}} \epsilon_{\mathbf{t}} f''\left(\mathbf{x}_{\mathbf{t}},\boldsymbol{\theta}\right)) + 2\sum_{\mathbf{t}\le\mathbf{n}} \left[f(\mathbf{x}_{\mathbf{t}},\boldsymbol{\theta}) - f(\mathbf{x}_{\mathbf{t}},\boldsymbol{\theta}_0)\right] f''(\mathbf{x}_{\mathbf{t}},\boldsymbol{\theta}) \\
& +2\sum_{\mathbf{t}\le\mathbf{n}} \left[f'\left(\mathbf{x}_{\mathbf{t}},\boldsymbol{\theta}\right)\right]\left[f'(\mathbf{x}_{\mathbf{t}},\boldsymbol{\theta})\right]^T .
\end{aligned}
$$

Now, we impose the following assumptions on the function $f(\cdot,\cdot)$ in addition to the Assumptions 1 – 3.

Assumption 4: $f(\mathbf{x}_{\mathbf{t}},\cdot)$ is twice continuously differentiable in $\boldsymbol{\Theta}$.

Assumption 5: Let $\{D_{\mathbf{n}},\ \mathbf{n}\in\mathbb{N}^k\}$ be a field of $k\times k$ non-singular matrices such that

(*i*) $\dfrac{1}{|\mathbf{n}|} D_{\mathbf{n}}^T \sum_{\mathbf{t}\le\mathbf{n}} \left[f'(\mathbf{x}_{\mathbf{t}},\boldsymbol{\theta})\right]\left[f'\left(\mathbf{x}_{\mathbf{t}},\boldsymbol{\theta}\right)\right]^T D_{\mathbf{n}}$ converges to a positive definite matrix $\Sigma(\boldsymbol{\theta}_0)$ uniformly as $|\mathbf{n}|\to\infty$ and $|\boldsymbol{\theta}-\boldsymbol{\theta}_0|\to 0$.

(*ii*) $\dfrac{1}{|\mathbf{n}|} D_{\mathbf{n}}^T \sum_{\mathbf{t}\le\mathbf{n}} \left[f(\mathbf{x}_{\mathbf{t}},\boldsymbol{\theta}) - f(\mathbf{x}_{\mathbf{t}},\boldsymbol{\theta})\right] f''\left(\mathbf{x}_{\mathbf{t}},\boldsymbol{\theta}\right) D_{\mathbf{n}} \longrightarrow 0$ uniformly as $|\mathbf{n}|\to\infty$ and $|\boldsymbol{\theta}-\boldsymbol{\theta}_0|\to 0$,

(*iii*) $\left\|D_{\mathbf{n}}^T f''(\mathbf{x}_{\mathbf{t}},\boldsymbol{\theta})\, D_{\mathbf{n}}\right\|_E \le M$ for all $\boldsymbol{\theta}\in\boldsymbol{\Theta}$. Here $\|\cdot\|_E$ is the Euclidean norm on matrices.

(*iv*) $\left\|D_{\mathbf{n}}^T(f''(\mathbf{x}_{\mathbf{t}},\boldsymbol{\theta}_1) - f''(\mathbf{x}_{\mathbf{t}},\boldsymbol{\theta}_2))D_{\mathbf{n}}\right\|_E \le b_{\mathbf{t},\mathbf{n}}|\boldsymbol{\theta}_1-\boldsymbol{\theta}_2|$ for all $\boldsymbol{\theta}_1\ne\boldsymbol{\theta}_2$,

where $b_{\mathbf{t},\mathbf{n}}>0, \mathbf{t}\in\mathbb{N}^k$ are some constants such that $\sum_{\mathbf{t}\le\mathbf{n}} b_{\mathbf{t},\mathbf{n}} = o(|\mathbf{n}|^3)$.

(*v*) $\max\limits_{1\le\mathbf{t}\le\mathbf{n}} \dfrac{1}{|\mathbf{n}|}\left\|D_{\mathbf{n}}^T f'(\mathbf{x}_{\mathbf{t}},\boldsymbol{\theta}_0)\right\|_E \to 0$ as $|\mathbf{n}|\to\infty$.

Theorem 3.1. *Under Assumptions 1 – 5 and (3),*

$$
\sqrt{|\mathbf{n}|}\,(D_{\mathbf{n}})^{-1}\left(\widehat{\boldsymbol{\theta}}_{\mathbf{n}}-\boldsymbol{\theta}_0\right) \xrightarrow{\mathcal{L}} N_k\left(\mathbf{0},\sigma^2\,\Sigma^{-1}\left(\boldsymbol{\theta}_0\right)\right) \ \text{as } |\mathbf{n}|\to\infty.
$$

Proof: From (24)

$$\sqrt{|\mathbf{n}|}\, D_{\mathbf{n}}^{-1}\, \left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right) = -\left(\frac{1}{|\mathbf{n}|}\, D_{\mathbf{n}}^T\, Q_{\mathbf{n}}''\, \left(\widehat{\boldsymbol{\theta}}_{*\mathbf{n}}\right)\, D_{\mathbf{n}}\right)^{-1} \frac{1}{\sqrt{|\mathbf{n}|}}\, D_{\mathbf{n}}^T\, Q_{\mathbf{n}}'(\boldsymbol{\theta}_0). \tag{27}$$

Now, from (26),

$$\begin{aligned} \frac{1}{\sqrt{|\mathbf{n}|}}\, D_{\mathbf{n}}^T\, Q_{\mathbf{n}}''\, (\boldsymbol{\theta})\, D_{\mathbf{n}} \;\; = \;\; & -2\, \frac{1}{|\mathbf{n}|} \sum_{\mathbf{t} \leq \mathbf{n}} \epsilon_{\mathbf{t}}\, \left[D_{\mathbf{n}}^T\, f''\, (\mathbf{x}_{\mathbf{t}}, \boldsymbol{\theta})\, D_{\mathbf{n}}\right] \\ & +2\, \frac{1}{|\mathbf{n}|}\, D_{\mathbf{n}}^T \sum_{\mathbf{t} \leq \mathbf{n}} \left[f\, (\mathbf{x}_{\mathbf{t}}, \boldsymbol{\theta}) - f\, (\mathbf{x}_{\mathbf{t}}, \boldsymbol{\theta}_0)\right]\, f''\, (\mathbf{x}_{\mathbf{t}}, \boldsymbol{\theta})\, D_{\mathbf{n}} \\ & + \frac{2}{|\mathbf{n}|}\, D_{\mathbf{n}}^T \sum_{\mathbf{t} \leq \mathbf{n}} \left[f'(\mathbf{x}_{\mathbf{t}}, \boldsymbol{\theta})\right]\, \left[f'(\mathbf{x}_{\mathbf{t}}, \boldsymbol{\theta})\right]^T\, D_{\mathbf{n}}. \end{aligned} \tag{28}$$

Using Assumption 5 parts (iii) and (iv), it can be shown as in the proof of Lemma 2.3 that

$$\sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \left\| \frac{1}{|\mathbf{n}|} \sum_{\mathbf{t} \leq \mathbf{n}} \epsilon_{\mathbf{t}}\, D_{\mathbf{n}}^T\, f''(\mathbf{x}_{\mathbf{t}}, \boldsymbol{\theta})\, D_{\mathbf{n}} \right\|_E \longrightarrow 0 \quad \text{in probability} \quad \text{as } |\mathbf{n}| \to \infty. \tag{29}$$

Since $\widehat{\boldsymbol{\theta}}_{\mathbf{n}} \to \boldsymbol{\theta}_0$ a.s., $\widehat{\boldsymbol{\theta}}_{*\mathbf{n}} \to \boldsymbol{\theta}_0$ a.s. as $|\mathbf{n}| \to \infty$, thus from (33), (34), and Assumption 5 parts (i) and (ii), we obtain

$$\frac{1}{|\mathbf{n}|}\, D_{\mathbf{n}}^T\, Q_{\mathbf{n}}''\, \left(\widehat{\boldsymbol{\theta}}_{*\mathbf{n}}\right)\, D_{\mathbf{n}} \longrightarrow 2\Sigma(\boldsymbol{\theta}_0) \quad \text{a.s. as} \quad |\mathbf{n}| \to \infty. \tag{30}$$

From (25)

$$\frac{1}{\sqrt{|\mathbf{n}|}}\, D_{\mathbf{n}}^T\, Q_{\mathbf{n}}'\, (\boldsymbol{\theta}_0) = -\frac{2}{\sqrt{|\mathbf{n}|}} \sum_{\mathbf{t} \leq \mathbf{n}} \epsilon_{\mathbf{t}}\, D_{\mathbf{n}}^T\, f'(\mathbf{x}_{\mathbf{t}}, \boldsymbol{\theta}_0). \tag{31}$$

To establish the asymptotic normality of the above, we first state the multidimension indices version of the Hajek-Sidek Theorem (Sen and Singer (1993)).

Lemma 3.2. *Let* $\{X_{\mathbf{n}},\ \mathbf{n} \in \mathbb{N}^k\}$ *be a field of independent and identically distributed random variables with mean* μ *and variance* σ^2*, and let* $\{c_{\mathbf{m},\mathbf{n}},\ \mathbf{m} \leq \mathbf{n} \in \mathbb{N}^k\}$ *be a field of* *real numbers such that*

$$\max_{1 \le \mathbf{m} \le \mathbf{n}} \frac{c^2_{\mathbf{m},\mathbf{n}}}{\sum\limits_{\mathbf{m} \le \mathbf{n}} c^2_{\mathbf{m},\mathbf{n}}} \longrightarrow 0 \quad \textit{as } |\mathbf{n}| \to \infty.$$

Then

$$\frac{\sum\limits_{\mathbf{m} \le \mathbf{n}} c^2_{\mathbf{m},\mathbf{n}} \, (X_{\mathbf{m}} - \mu)}{\sqrt{\sum\limits_{\mathbf{m} \le \mathbf{n}} c^2_{\mathbf{m},\mathbf{n}}}} \stackrel{\mathcal{L}}{\longrightarrow} N(0, \sigma^2) \quad \textit{as } |\mathbf{n}| \to \infty.$$

The proof of this follows along the standard lines of Hajeck-Sidek Theorem.

Using this Lemma, and Assumption 5 parts (i) and (iv), we have, for any $\boldsymbol{\lambda} \in \mathbb{R}^k$,

$$\frac{\boldsymbol{\lambda}^T \sum\limits_{\mathbf{t} \le \mathbf{n}} \epsilon_{\mathbf{t}} \, D^T_{\mathbf{n}} \, f'(\mathbf{x}_{\mathbf{t}}, \boldsymbol{\theta}_0)}{\sqrt{\boldsymbol{\lambda}^T \sum\limits_{\mathbf{t} \le \mathbf{n}} D^T_{\mathbf{n}} \, [f'(\mathbf{x}_{\mathbf{t}}, \boldsymbol{\theta}_0)] \, [f'(\mathbf{x}_{\mathbf{t}}, \boldsymbol{\theta}_0)]^T \, D_{\mathbf{n}} \, \boldsymbol{\lambda}}} \stackrel{\mathcal{L}}{\longrightarrow} N(0, \sigma^2).$$

From Assumption 5 (i), we obtain

$$\boldsymbol{\lambda}^T \frac{1}{\sqrt{|\mathbf{n}|}} \sum_{\mathbf{t} \le \mathbf{n}} \epsilon_{\mathbf{t}} \, D^T_{\mathbf{n}} \, f'(\mathbf{x}_{\mathbf{t}}, \boldsymbol{\theta}_0) \stackrel{\mathcal{L}}{\longrightarrow} N\left(0, \sigma^2 \, \boldsymbol{\lambda}^T \, \Sigma(\boldsymbol{\theta}_0) \, \boldsymbol{\lambda}\right) \text{ for any } \boldsymbol{\lambda} \in \mathbb{R}^k,$$

and in view of (29),

$$\frac{1}{\sqrt{|\mathbf{n}|}} D^T_{\mathbf{n}} \, Q'_{\mathbf{n}} \, (\boldsymbol{\theta}_0) \stackrel{\mathcal{L}}{\longrightarrow} N\left(\mathbf{0}, 4\sigma^2 \Sigma(\boldsymbol{\theta}_0)\right).$$

Now, the proof of the theorem follows from (27) and (30).

Remark. If Assumption 5 holds only for a subsequence of $\{\mathbf{n}, \ \mathbf{n} \in \mathbb{N}^k\}$ such that $|\mathbf{n}| \to \infty$, then the result of Theorem 3.1 holds for that subsequence. This will be seen in the example below.

We now consider asymptotic normality of the least squares estimators for model (5). We will show that the parameters of the asymptotic normal distribution depend on subsequences of $\mathbf{n}$. Asymptotic normality can be obtained for two types of sequences: (i) a subsequence for which $\min(n_1, n_2) \to \infty$, and (ii) a subsequence for which any one of n_1 and $n_2 \to \infty$ while the other kept constant. For notational convenience, as in Section 2, we assume that $p = 1$ and deal only with model (22). Let $\widehat{\boldsymbol{\theta}}_{\mathbf{n}} = \left(\widehat{\alpha}_{\mathbf{n}}, \widehat{\lambda}_{1\mathbf{n}}, \widehat{\lambda}_{2\mathbf{n}}\right)^T$ be the least squares estimators, and $\boldsymbol{\theta}_0 = (\alpha_0, \lambda_{10}, \lambda_{20})^T$ be the true parameter vector. Let

$$D_{\mathbf{n}} = \text{Diag } \left(1, n_1^{-1}, n_2^{-1}\right).$$

In order to apply Theorem 3.1, we need to verify Assumptions 4 and 5. Assumption 4 is clearly satisfied. To verify Assumption 5, observe that

$$f'(\mathbf{x_t}, \boldsymbol{\theta}) = \begin{bmatrix} \cos(t_1\lambda_1 + t_2\lambda_2) \\ -\alpha\, t_1 \sin(t_1\lambda_1 + t_2\lambda_2) \\ -\alpha\, t_2 \sin(t_1\lambda_1 + t_2\lambda_2) \end{bmatrix}. \tag{32}$$

Thus

$$\sum_{\mathbf{t}\leq\mathbf{n}} \left[f'\left(\mathbf{x_t}, \boldsymbol{\theta}\right)\right] \left[f'\left(\mathbf{x_t}, \boldsymbol{\theta}\right)\right]^T = \tag{33}$$

$$\begin{bmatrix} \sum_{\mathbf{t}\leq\mathbf{n}} \cos^2(t_1\lambda_1 + t_2\lambda_2) & -\frac{\alpha}{2} \sum_{\mathbf{t}\leq\mathbf{n}} t_1 \sin 2(t_1\lambda_1 + t_2\lambda_2) & -\frac{\alpha}{2} \sum_{\mathbf{t}\leq\mathbf{n}} t_2 \sin 2(t_1\lambda_1 + t_2\lambda_2) \\ -\frac{\alpha}{2} \sum_{\mathbf{t}\leq\mathbf{n}} t_1 \sin 2(t_1\lambda_1 + t_2\lambda_2) & \alpha^2 \sum_{\mathbf{t}\leq\mathbf{n}} t_1^2 \sin^2(t_1\lambda_1 + t_2\lambda_2) & \alpha^2 \sum_{\mathbf{t}\leq\mathbf{n}} t_1 t_2 \sin^2(t_1\lambda_1 + t_2\lambda_2) \\ -\frac{\alpha}{2} \sum_{\mathbf{t}\leq\mathbf{n}} t_2 \sin 2(t_1\lambda_1 + t_2\lambda_2) & \alpha^2 \sum_{\mathbf{t}\leq\mathbf{n}} t_1 t_2 \sin^2(t_1\lambda_1 + t_2\lambda_2) & \alpha^2 \sum_{\mathbf{t}\leq\mathbf{n}} t_2^2 \sin^2(t_1\lambda_1 + t_2\lambda_2) \end{bmatrix}.$$

Note that

$$\sum_{\mathbf{t}\le\mathbf{n}} \cos 2\left(t_1\lambda_1+t_2\lambda_2\right) = \frac{1}{2}\sum_{\mathbf{t}\le\mathbf{n}}\left[e^{2i(t_1\lambda_1+t_2\lambda_2)}+e^{-2i(t_1\lambda_1+t_2\lambda_2)}\right] \tag{34}$$

$$= \frac{1}{2}\left[\sum_{1\le t_1\le n_1} e^{2i\lambda_1 t_1}\sum_{1\le t_2\le n_2} e^{2i\lambda_2 t_2}+\sum_{1\le t_1\le n_1} e^{-2i\lambda_1 t_1}\sum_{1\le t_2\le n_2} e^{-2i\lambda_2 t_2}\right]$$

$$= \frac{1}{2}\left[e^{2i(\lambda_1+\lambda_2)}\frac{\left(1-e^{2i\lambda_1 n_1}\right)\left(1-e^{2i\lambda_2 n_2}\right)}{\left(1-e^{2i\lambda_1}\right)\left(1-e^{2i\lambda_2}\right)}+e^{-2i(\lambda_1+\lambda_2)}\frac{\left(1-e^{-2i\lambda_1 n_1}\right)\left(1-e^{-2i\lambda_1 n_1}\right)}{\left(1-e^{-2i\lambda_1}\right)\left(1-e^{-2i\lambda_2}\right)}\right].$$

This term is of order $O(1)$. This implies that

$$\frac{1}{|\mathbf{n}|}\sum_{\mathbf{t}\le\mathbf{n}}\cos^2(t_1\lambda_1+t_2\lambda_2) = \frac{1}{2}+\frac{1}{2|\mathbf{n}|}\sum_{\mathbf{t}\le\mathbf{n}}\cos 2\left(t_1\lambda_1+t_2\lambda_2\right)\to\frac{1}{2} \text{ as } |\mathbf{n}|\to\infty.$$

By differentiating (34) with respect to λ_1, we get

$$\sum_{\mathbf{t}\le\mathbf{n}} t_1 \sin 2(t_1\lambda_1+t_2\lambda_2) = O(n_1).$$

Thus

$$\frac{1}{|\mathbf{n}|}\frac{1}{n_1}\sum_{\mathbf{t}\le\mathbf{n}} t_1 \sin 2(t_1\lambda_1+t_2\lambda_2)\longrightarrow 0 \quad \text{as} \quad |\mathbf{n}|\to\infty.$$

Similarly, it can be seen that

$$\frac{1}{|\mathbf{n}|}\frac{1}{n_2}\sum_{\mathbf{t}\le\mathbf{n}} t_2 \sin^2(t_1\lambda_1+t_2\lambda_2)\longrightarrow 0 \quad \text{as} \quad |\mathbf{n}|\to\infty.$$

By differentiating (34) twice with respect to λ_1, we get

$$\sum_{\mathbf{t}\le\mathbf{n}} t_1^2 \cos 2(t_1\lambda_1+t_2\lambda_2) = O(n_1^2).$$

Thus

$$\begin{aligned}
\frac{1}{|\mathbf{n}|}\frac{1}{n_1^2}\sum_{\mathbf{t}\le\mathbf{n}} t_1 \sin^2(t_1\lambda_1+t_2\lambda_2) &= \frac{1}{2|\mathbf{n}|}\frac{1}{n_1^2}\sum_{\mathbf{t}\le\mathbf{n}} t_1^2 - \frac{1}{2|\mathbf{n}|}\frac{1}{n_1^2}\sum_{\mathbf{t}\le\mathbf{n}} t_1^2 \cos 2(t_1\lambda_1+t_2\lambda_2)\\
&= \frac{(n_1+1)(2n_1+1)}{12\, n_1^2} - \frac{1}{2|\mathbf{n}|}\frac{1}{n_1^2}\sum_{\mathbf{t}\le\mathbf{n}} t_1^2 \cos^2(t_1\lambda_1+t_2\lambda_2),
\end{aligned}$$

converges to $1/6$ if $\min(n_1,n_2)\to\infty$ or if $n_1\to\infty$, and converges to $(n_1+1)(2n_1+1)/12n_1^2$ if n_1 is fixed but $n_2\to\infty$.

Similarly, it can be seen that $\frac{1}{|\mathbf{n}|}\frac{1}{n_1n_2}\sum_{\mathbf{t}\le\mathbf{n}} t_1t_2\sin^2(t_1\lambda_1+t_2\lambda_2)$ converges to $1/8$ if $\min(n_1,n_2)\to\infty$, and converges to $(n_1+1)/8n_1$ if n_1 is fixed but $n_2\to\infty$, and converges to $(n_2+1)/8n_2$ if n_2 is fixed but $n_1\to\infty$. Also $\frac{1}{|\mathbf{n}|}\frac{1}{n_2}\sum_{\mathbf{t}\le\mathbf{n}} t_2^2\sin^2 2(t_1\lambda_1+t_2\lambda_2)$ converges to $1/6$ if $\min(n_1,n_2)\to\infty$ or if $n_2\to\infty$, and converges to $(n_2+1)(2n_2+1)/12n_2^2$ if n_2 is fixed but $n_1\to\infty$.

Thus, we have that Assumption 5 (i) is satisfied with

$$\Sigma(\boldsymbol{\theta}_0)=\begin{bmatrix}\frac{1}{2} & 0 & 0\\ 0 & \alpha_0^2/6 & \alpha_0^2/8\\ 0 & \alpha_0^2/8 & \alpha_0^2/6\end{bmatrix} \quad \text{if} \quad \min(n_1,n_2)\to\infty, \tag{35}$$

with

$$\Sigma(\boldsymbol{\theta}_0)=\begin{bmatrix}\frac{1}{2} & 0 & 0\\ 0 & \alpha_0^2\frac{(n_1+1)(2n_1+1)}{12n_1^2} & \alpha_0^2\frac{(n_1+1)}{8n_1}\\ 0 & \alpha_0^2\frac{(n_1+1)}{8n_1} & \alpha_0^2/6\end{bmatrix} \quad \text{if} \quad n_1 \text{ is fixed but} \quad n_2\to\infty, \tag{36}$$

and with

$$\Sigma(\boldsymbol{\theta}_0) = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \alpha_0^2/6 & \alpha_0^2 \frac{(n_2+1)}{8n_1} \\ 0 & \alpha_0^2 \frac{(n_2+1)}{8n_2} & \alpha_0^2 \frac{(n_2+1)(2n_2+1)}{12n_2^2} \end{bmatrix} \quad \text{if} \quad n_2 \text{ is fixed but} \quad n_1 \to \infty. \tag{37}$$

It is easy to see that Assumption 5 parts (iii) and (iv) are satisfied. To verify Assumption 5 (ii), note that

$$\frac{1}{|\mathbf{n}|} D_{\mathbf{n}}^T \sum_{\mathbf{t} \leq \mathbf{n}} [f(\mathbf{x_t}, \boldsymbol{\theta}) - f(\mathbf{x_t}, \boldsymbol{\theta}_0)] \; f''(\mathbf{x_t}, \boldsymbol{\theta}_0) \, D_{\mathbf{n}} = \frac{1}{|\mathbf{n}|} D_{\mathbf{n}}^T \sum_{\mathbf{t} \leq \mathbf{n}} [f'(\mathbf{x_t}, \boldsymbol{\theta}_{t*})]^T \; (\boldsymbol{\theta} - \boldsymbol{\theta}_0) \, f''(\mathbf{x_t}, \boldsymbol{\theta}) \, D_{\mathbf{n}},$$

where $\boldsymbol{\theta}_{\mathbf{t}*} = h_t \boldsymbol{\theta}_0 + (1 - h_t)\boldsymbol{\theta}$ for some $0 \leq h_t \leq 1$.

Assumption 5 (ii) now can be verified from (32) and Assumption 5 (iii).

Thus, from the remark after the proof of Theorem 3.1, we conclude that

$$\sqrt{|\mathbf{n}|} \; \left((\widehat{\alpha} - \alpha_0) \, , \; n_1 \left(\widehat{\lambda}_1 - \lambda_{10} \right) \, , \; n_2 \left(\widehat{\lambda}_2 - \lambda_{20} \right) \right) \xrightarrow{\mathcal{L}} \; N_3 \left(\mathbf{0}, \sigma^2 \, \Sigma^{-1}(\boldsymbol{\theta}_0) \right)$$

as either $\min(n_1, n_2) \to \infty$, or $n_2 \to \infty$ while n_1 is held fixed, or $n_1 \to \infty$ while n_2 is held fixed, where $\Sigma(\boldsymbol{\theta}_0)$ is given by (35) – (37) for the appropriate subsequences.

Acknowledgement:

The authors are grateful to Jim Kuelbs for his helpful comments and suggestions. We are indebted to Hira Koul for his careful reading of this manuscript and his invaluable suggestions resulting in a correction and improvement of the presentation of the content.

References:

Francos, J. M., Meiri, A. Z. and Porat, B. (1993). A united texture model based on a 2-D Wold like decomposition. *IEEE Trans. Signal Processing*, **41**, 2665–2678.

Gallant, A. R. (1987). *Nonlinear Statistical Models*. J. Wiley.

Francos, J. M., Meiri, A. Z. and Porat, B. (1993). A united texture

Jennrich, R. I. (1969). Asymptotic properties of non-linear least squares estimators. *Ann. Math. Statist.*, **40**, 633–643.

Kundu, D. (1993). Asymptotic theory of least squares estimator of a particular non-linear regression model. *Statistics & Probability Letters*, **18**, 13–17.

Mandrekar, V. and Zhang, H. (1996). On parameter estimation in a texture model. Submitted for publication.

McClellan, J. H. (1982). Multidimensional spectral estimation. *Proc. IEEE*, **70**, 1029–1039.

Mikosch, T. and Norvaisa, R. (1987). Strong laws of large numbers for fields of Banach space valued random variables. *Prob. Th. Rel. Fields*, **74**, 241–253.

Rao, C. R., Zhao, L. and Zhou, B. (1994). Maximum likelihood estimation of 2-D superimpared exponential signals. *IEEE Trans. Signal Processing*, **42**, 1795–1802.

Sen, P. K. and Singer, J. M. (1993). *Large Sample Methods in Statistics: An Introduction with Applications.* Chapman & Hall.

Wu, C. F. (1981). Asymptotic theory of non-linear least square estimation, *Ann. Statist.*, **9**, 501–513.

Yuan, T. and Subba Rao, T. (1993). Spectrum estimation for random fields with applications to Markov modelling and texture classification, in *Markov Random Fields, Theory and Applications*, edited by R. Chellappa and A. K. Jain, Academic Press.