MAXIMUM LIKELIHOOD ESTIMATION OF LOGISTIC SINUSOIDAL REGRESSION MODELS

Yu Weng

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

December 2013

APPROVED:

Kai-Sheng Song, Major Professor
Pieter Allaart, Committee Member
Jiaping Wang, Committee Member
Su Gao, Chair of the Department of Mathematics
Mark Wardell, Dean of the Toulouse Graduate School

We consider the problem of maximum likelihood estimation of logistic sinusoidal regression models and develop some asymptotic theory including the consistency and joint rates of convergence for the maximum likelihood estimators. The key techniques build upon a synthesis of the results of Walker and Song and Li for the widely studied sinusoidal regression model and on making a connection to a result of Radchenko. Monte Carlo simulations are also presented to demonstrate the finite-sample performance of the estimators.
ACKNOWLEDGMENTS

Heartfelt gratitude and appreciation go to my advisor Dr. Kai-Sheng Song for his dedicated guidance, advice, and patience during the dissertation process. I am immensely fortunate to have him as an advisor and I enjoyed studying under him. He shared both the subject knowledge and occasionally a laugh with me, and without him this thesis would not have come to its completion. Thanks also go to Dr. Pieter Allaart and Dr. Jiaping Wang for their helpful comments on the presentation. I thank the UNT Department of Mathematics for financial support and faculty mentoring I received during my graduate study, and for the opportunity to serve by teaching, grading, and tutoring. I have benefited from Dr. Mauldin and Dr. Neuberger's kind words and cheerful devotion to their students. I am fortunate to have had Carol Fichera and Chelsea Hansard Solano as my supervisors, who have trained me professionally and cheerfully, making it a pleasure and fulfillment to be at work.

Last but not least, thanks go to my friends, family, and extended families for the many meaningful discussions and time together, which kept the work and study in a broader perspective. Thanks are long overdue to my parents Tongsen Weng and Meiying Chen for their patient support and love. Special thanks go to my excellent friends Tammy Lim, Grace Lee, and Bonita Ng, as well as many others, for their precious friendships and help throughout the process. Besides numerous others, I owe my growth as a person tremendously to my pastor, Pastor Teo, who is an example of what it means to give to people and to encourage others, whether in health or in sickness, in prosperity or in poverty, in happiness or in suffering.
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.4. Proof of Corollary 2</td>
<td>33</td>
</tr>
<tr>
<td>1.5. Proof of Lemma 4</td>
<td>33</td>
</tr>
<tr>
<td>1.6. Proof of Modified Lemma 5</td>
<td>34</td>
</tr>
<tr>
<td>1.7. Proof of Lemma 6</td>
<td>35</td>
</tr>
<tr>
<td>1.8. Proof of Lemma 7</td>
<td>36</td>
</tr>
<tr>
<td>2. Radchenko’s Lemma 1</td>
<td>45</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>47</td>
</tr>
</tbody>
</table>
LIST OF TABLES

3.1 Variance vs. Time series length 28
3.2 MSE vs. Time series length 28
3.3 Bias vs. Time series length 28
3.4 Means vs. Time series length 28
3.5 KS test for normality: p-values 29
3.6 Shapiro-Wilks test: p-values 29
LIST OF FIGURES

3.1 Root mean square error vs. Length of time series  25
3.2 Bias vs. Length of time series  26
3.3 Variance of estimates vs. Length of time series  27
CHAPTER 1

LITERATURE REVIEW AND INTRODUCTION TO THE PROBLEM

1.1. Known Results and Techniques on Linear and Nonlinear Regression Models

The sinusoidal regression model for a time series \{y_1, \ldots, y_n\} given by

\[ y_t = \sum_{k=1}^{l} \{A_{0k} \cos(\omega_{0k} t) + B_{0k} \sin(\omega_{0k} t)\} + \epsilon_t \]

for \( t = 1, \ldots, n \) under an assumption of i.i.d. white noise \( \epsilon_t \), real-valued amplitudes \( A_{0k}, B_{0k} \), and a frequency \( \omega_{0k} \in (0, \pi) \) for all \( k \in \{1, \ldots, l\} \) has applications in a wide range of fields including bioinformatics, engineering, and physical sciences. The problem of estimating the parameter \( \theta_0 := [A_{01}, B_{01}, \omega_{01}, \ldots, A_{0l}, B_{0l}, \omega_{0l}]' \) of the sinusoidal regression model belongs to the larger class of nonlinear regression problems. In this paper, we will review some results by Wu, Song & Li, and Radchenko, and consider a closely related logistic sinusoidal regression model and derive the consistency and joint rate of convergence results of the maximum likelihood estimator of the parameter \( [A_0, B_0, \omega_0] \) in the model given by independent \( y_1, \ldots y_n \), \( y_t \sim \text{Bernoulli}(p_{0t}) \), and

\[ \log \frac{p_{0t}}{1 - p_{0t}} = A_0 \cos \omega_0 t + B_0 \sin \omega_0 t \]

We will start by reviewing some results on nonlinear regression. Several sufficient conditions for the consistency and asymptotic normality for the least squares estimator have been sought for and provided, for examples in Wu [12] and Jennrich [2]. Wu noted that due to the nonlinear nature of the nonlinear least squares estimation problem, the estimator lacks the finite sample optimality properties, and hence the general theory developed surrounding the nonlinear least squares estimation problems is asymptotic. Wu[12] and Jennrich [2] independently established sufficient conditions in which asymptotic consistency and asymptotic normality can be obtained in a nonlinear least squares estimation procedure; the latter author included a proof of the existence of a measurable nonlinear least squares estimator in his
Lemma 2 under fairly general regularity assumptions on the parameter space, the criterion function, and the independent, identically distributed finite variance residuals.

While the results of Jennrich and Wu developed contrasting sets of sufficient conditions, it is worth noting that their conditions make different assumptions on growth rate behavior on the original function, the gradient, and the Hessian. With specific examples, Wu compared his results with those of Jennrich’s in his own paper.

The assumptions on which the asymptotic normality is derived as in Wu require that the components of the parameter vector $\hat{\theta}_n$ converge to the true parameter $\theta_0$ at the same rates. With this specific configuration, Wu’s results cannot be automatically extended where the rates of the components are different.

For the sinusoidal regression problem, Walker proved the consistency of the parameter estimate $(\hat{A}_n, \hat{B}_n, \hat{\omega}_n)$ in a periodogram approximation to the criterion function $S_n(\theta_n) := \sum_{t=1}^n (y_t - (A_n \cos \omega_n t + B_n \sin \omega_n t))^2$. However, it was noted in Song & Li’s paper that his result did not automatically result in a proof that the nonlinear least squares estimator based on the original criterion function $S_n(\theta_n)$ is consistent. On the other hand, even if one is to prove that the NLS estimator based on the original criterion function is consistent using Wu’s approach, his result is not directly applicable. Song & Li note an important reason why the result in Wu cannot be directly applied; namely, the assumption of uniform rates on the components of the parameter being estimated in Wu’s Theorems 2 and 3 is not automatically true in the setting of the sinusoidal regression problem.

However, Song & Li did obtain a condition based on Wu’s Lemma 1 which would be sufficient to deduce consistency of all components of the parameter simultaneously.

To use Wu’s results on a model, one needs to check if the model satisfies the assumptions upon which consistency is derived. Song & Li points out that Wu’s result is not directly applicable to the sinusoidal regression problem. The main difficulty stems from the fact that Wu’s condition when specialized to the sinusoidal regression problem depends on the proof of the existence of a positive lower bound for the liminf of $\inf_{\theta \in \Theta_\delta} \frac{1}{n} \sum_{t=1}^n (f_t(\theta) - f_t(\theta_0))^2$ whenever $\delta > 0$ and $\Theta_\delta := \{ \theta \in \Theta : \| \theta - \theta_0 \| > \delta \}$ is given.
Instead of attempting to prove consistency of the parameter based on Wu’s Lemma 1 and a sufficient formulation in the sinusoidal regression model, which when successful would mean a simultaneous consistency result for all components of the parameter, Song & Li followed a different line of argument: weak consistency of the frequency parameter is established first by recognizing the equivalence between nonlinear least squares and residual sum of squares maximization and utilizing Walker’s proof of consistency for the frequency component in a periodogram approximation to the original problem of nonlinear least squares. With some application of linear algebra, the weak consistency of the amplitude components is proved with a use of An, Chen & Hannan’s [1] asymptotic result that the periodogram is \( O(\sqrt{n\log n}) \) almost surely.

Song & Li proceeded to prove the asymptotic normality of the parameter \( \hat{\theta}_n \) scaled componentwise by their respective rates. The proof is accomplished by showing the convergence in probability of the components of the matrix \( D_n \nabla^2 S_n(\hat{\theta}) D_n \) uniformly in \( \hat{\theta} \) to a positive-definite matrix \( U \), where \( D_n \) is the diagonal square matrix \( \text{diag}(n^{-1/2}, n^{-1/2}, n^{-3/2}) \), with \( S_n \) representing the sum of squared differences between the observed value \( y_t \) and the model evaluated at \( \tilde{\theta} \), \( t \) ranging from 1 to \( n \). The uniform convergence relies on consistency and Walker’s [11] result that

\[
\left( n^{-(r+1)} \sum_{t=1}^{n} t^r \exp(it\tilde{\omega}_k) \epsilon_t \right) = O_p(n^{-1/4}) \quad \text{uniformly in } \tilde{\omega}_k.
\]

1.2. The Logistic Regression Model and Results

The main subject in this paper is the logistic sinusoidal regression model. In this paper, the proof of consistency of the maximum likelihood estimator in the logistic sinusoidal regression model will be based on the assumptions that the components of the parameter lie in a compact set, with the additional requirement that the frequency component is in a compact set between 0 and \( \pi \) and bounded away from 0 and \( \pi \). The problem of estimating the parameter hence becomes finding a MLE of the model. This is the same as estimating the minimizer of the negative of the log-likelihood function. To ease the borrowing of notation and results surrounding least squares estimation from Wu, Song & Li, and Walker, we will choose to minimize \( G_n \) in this paper, where \( G_n \) represents the negative of the log-likelihood function.
Proof of consistency will be pursued by first using Taylor expansion and mean value theory to rewrite the original negative maximum likelihood equation. Motivated by Wu’s Lemma 1 and Song & Li’s discussions, a sufficient condition for consistency of the parameter in the logistic regression model is derived.

In particular, several inequalities and measure theoretic results from Walker [11] are borrowed, while Song & Li’s matrix representation and techniques are adopted and catered to the specifics of this problem which result in the joint consistency of the parameter \( \{ \hat{A}_n, \hat{B}_n, \hat{\omega}_n \} \). Note that the proof discussed in the subsequent chapter breaks down the sufficient condition for consistency in two separate cases. In fact, the same consistency proof process reveals an alternative proof to proving asymptotic consistency of the parameter in the corresponding sinusoidal regression problem. It is worth noting that Walker and Song & Li’s results and methods are integral to the success of this consistency proof.

Next, the notation in Song & Li is reproduced into our proof of the parameter’s joint rate of convergence. For a succinct representation of the transition from a prerequisite of consistency to the proof of asymptotic normality of a point-estimator, we may take a look at [8]’s Chapter 2.1. While the rates of convergence for the sinusoidal model parameter is rather applied directly in the proof of asymptotic normality of a point-estimate in Song & Li based on a previously proved rates of convergence result by [9], we will explicitly identify the rates of convergence for the logistic regression model parameter with Radchenko’s Lemma 1 [7].

Specifically, to derive the rate of convergence using Radchenko’s Lemma 1, one needs to write the criterion function \( \frac{1}{n} \left\{ G_n(\hat{\theta}_n) - G_n(\theta_0) \right\} \) as a sum of a stochastic function and a ”positive” function (in the sense that it converges in probability to a positive quantity bounded below by an expression involving a positive linear combination of the positive powers of the norms of the parameter components). With a reparametrization, it is shown that the configuration for application of Radchenko’s Lemma 1 is satisfied and that the joint rate of convergence for the parameter \( \hat{\theta}_n \) is given by \( \text{diag}(1,1,n)(\hat{\theta}_n) = O_p(n^{-1/2}) \).
CHAPTER 2
LOGISTIC SINUSOIDAL REGRESSION

2.1. Introduction

Consider the problem of estimating the parameter \( \theta_0 := (A_0, B_0, \omega_0)' \) from a time series \( \{y_1, y_2, ..., y_n\} \), where the \( y_i \)'s are independent random variables modeled by \( Bernoulli(p_{0t}) \), and the probability parameters \( p_{0t} \) are linked to \( \theta_0 \) by the logit link function:

\[
\log \frac{p_{0t}}{1 - p_{0t}} = A_0 \cos \omega_0 t + B_0 \sin \omega_0 t
\]

where \( A_0, B_0 \in \mathbb{R} \) and \( \omega_0 \in (0, \pi) \) are unknown constants.

Note that from the above model it follows that

\[
p_{0t} = \frac{\exp(A_0 \cos \omega_0 t + B_0 \sin \omega_0 t)}{1 + \exp(A_0 \cos \omega_0 t + B_0 \sin \omega_0 t)}
\]

The maximum likelihood estimator of \( \theta_0 \) is alternatively given as the minimizer of the negative of the log likelihood function

\[
G_n(\theta) := -\sum_{t=1}^{n} \log l(y_t|\theta) = -\sum_{t=1}^{n} \{y_t \log \frac{p_t}{1 - p_t} + \log(1 - p_t)\}
\]

in which \( p_t := p_t(\theta) := \frac{\exp(A \cos \omega t + B \sin \omega t)}{1 + \exp(A \cos \omega t + B \sin \omega t)} \), where \( \theta := (A, B, \omega)' \), so \( p_{0t} \) is alternatively represented as \( p_t(\theta_0) \). Define \( f(\theta, t) := A \cos \omega t + B \sin \omega t \). Denote the minimizer of \( G_n(\theta) \) over some parameter set \( \Theta \subset \Theta_0 := \{\theta : A, B \in \mathbb{R}; 0 < \omega < \pi\} \) as \( \hat{\theta}_n \).

We will set out to prove:

(1) The consistency of \( \hat{\theta}_n \), but even stronger, that \( \hat{\theta}_n^* := \text{diag}(1, 1, n)\hat{\theta}_n = o_p(1) \).

(2) The joint rate of convergence of \( \hat{\theta}_n^* \).

The mean value theory for the Taylor expansion of \( G_n(\theta) \) at \( \theta = \theta_0 \) gives

\[
G_n(\theta) - G_n(\theta_0) = \nabla G_n(\theta_0)'(\theta - \theta_0) + (\theta - \theta_0)'H_n(\theta)(\theta - \theta_0)
\]

where

\[
H_n(\theta) := \int_{0}^{1} (1 - s)\nabla^2 G_n(\theta_0 + s(\theta - \theta_0)) \, ds;
\]
is the integral form of the remainder in the first order Taylor expansion.

\[ \nabla G_n(\theta_0) := \left( -\sum_{t=1}^{n} (y_t - p_{0t}) \cos \omega_0 t, -\sum_{t=1}^{n} (y_t - p_{0t}) \sin \omega_0 t, -\sum_{t=1}^{n} (y_t - p_{0t}) t (-A_0 \sin \omega_0 t + B_0 \cos \omega_0 t) \right) \]

with \( \nabla \) representing the gradient operator and \( \nabla^2 \) representing the Hessian operator. Note that since \( \hat{\theta}_n \) minimizes \( G_n(\theta) \), it follows that

\[ 0 \geq n^{-1} \nabla G_n(\theta_0)'(\hat{\theta}_n - \theta_0) + n^{-1}(\hat{\theta}_n - \theta_0)'H_n(\hat{\theta}_n)(\hat{\theta}_n - \theta_0) \]

\[ = n^{-1/2}(\hat{\theta}_n^*)'D_n \nabla G_n(\theta_0) + (\hat{\theta}_n^*)'D_n H_n(\hat{\theta}_n)D_n(\hat{\theta}_n^*) \]

where

\[ D_n := \text{diag}(n^{-1/2}, n^{-1/2}, n^{-3/2}) \]

Let \( U_n(\theta) := D_n \nabla^2 G_n(\theta) D_n \).

In the first section below, we will prove the consistency of \( \hat{\theta}_n^* \).

In the second section, we will show that based on the consistency, the following can be achieved:

\[ (1) \quad U_n(\theta_0) \rightarrow_p U \]

where

\[ U := \begin{bmatrix} d_1 & d_3 & \frac{1}{2}(-A_0 d_3 + B_0 d_1) \\ Symmetry & d_2 & \frac{1}{2}(-A_0 d_2 + B_0 d_3) \\ & & \frac{1}{3}(A_0^2 d_2 + B_0^2 d_1 - 2A_0 B_0 d_3) \end{bmatrix} \]

with

\[ d_1 := \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} p_{0t} (1 - p_{0t}) \cos^2(\omega_0 t) \]

\[ d_2 := \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} p_{0t} (1 - p_{0t}) \sin^2(\omega_0 t) \]

\[ d_3 := \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} p_{0t} (1 - p_{0t}) \cos(\omega_0 t) \sin(\omega_0 t) \]
and that uniformly in $\tilde{\theta}_n := \theta_0 + s(\hat{\theta}_n - \theta_0)$ for all $s \in [0, 1]$,

(2) \quad U_n(\tilde{\theta}_n) - U_n(\theta_0) \rightarrow P_0

where $U_n(\tilde{\theta}_n)$ is symmetric and

\begin{align*}
U_n(\tilde{\theta}_n)_{11} &= n^{-1} \tilde{p}_t(1 - \tilde{p}_t)(\cos \tilde{\omega}_n t)^2 \\
U_n(\tilde{\theta}_n)_{12} &= n^{-1} \sum_{t=1}^n \tilde{p}_t(1 - \tilde{p}_t)\cos \tilde{\omega}_n t \sin \tilde{\omega}_n t \\
U_n(\tilde{\theta}_n)_{13} &= n^{-2} \sum_{t=1}^n \tilde{p}_t(1 - \tilde{p}_t)(-A\sin \tilde{\omega}_n t + B\cos \tilde{\omega}_n t) t \cos \tilde{\omega}_n t + t(y_t - \tilde{p}_t) \sin \tilde{\omega}_n t \\
U_n(\tilde{\theta}_n)_{22} &= n^{-1} \sum_{t=1}^n \tilde{p}_t(1 - \tilde{p}_t)(\sin \tilde{\omega}_n t)^2 \\
U_n(\tilde{\theta}_n)_{23} &= n^{-2} \sum_{t=1}^n \tilde{p}_t(1 - \tilde{p}_t)(-A\sin \tilde{\omega}_n t + B\cos \tilde{\omega}_n t)t \sin \tilde{\omega}_n t - t(y_t - \tilde{p}_t) \cos \tilde{\omega}_n t \\
U_n(\tilde{\theta}_n)_{33} &= n^{-3} \sum_{t=1}^n \tilde{p}_t(1 - \tilde{p}_t)(-A\sin \tilde{\omega}_n t + B\cos \tilde{\omega}_n t)^2 t^2 + t^2(y_t - \tilde{p}_t)(A\cos \tilde{\omega}_n t + B\sin \tilde{\omega}_n t)
\end{align*}

The ability to conclude (2) is the basis for approximating $G_n(\theta) - G_n(\theta_0)$ using the Taylor series expansion of the second order with $D_nH_n(\hat{\theta}_n)D_n$ replaced by $\frac{1}{2}U_n(\theta_0)$. Note that we will need to show that $d_1, d_2,$ and $d_3$ are well defined, which we will accomplish based on two lemmas. The first of which, we call finite truncation, and the second, the trigonometric big oh identities.

We will begin by proving two key lemmas.

**Lemma 1.** Suppose $f : \mathbb{R} \rightarrow \mathbb{R}$ is real analytic on all of $\mathbb{R}$, and $X : \mathbb{N} \rightarrow \mathbb{R}$ is a bounded function. Then $\forall k \in \mathbb{N} \cup \{0\}$, and $\forall \delta > 0$, there exists $M(\delta) \in \mathbb{N}$ and $\{a_0, \ldots, a_{M(\delta)}\}$ such that

$$\left| \frac{1}{n^{k+1}} \sum_{t=1}^n t^k f(X(t)) - \frac{1}{n^{k+1}} \sum_{t=1}^n t^k \sum_{m=0}^{M(\delta)} a_m X(t)^m \right| < \delta$$

7
Lemma 2. For all $j, k \in \mathbb{N} \cup \{0\}, \omega, \omega_0 \in \mathbb{R}$,

$$P(j, k, \omega) := \sum_{t=1}^{n} (\cos \omega_0 t)^j \cos \omega t (\sin \omega_0 t)^k = b_{jk} n + O(1)$$

$$Q(j, k, \omega) := \sum_{t=1}^{n} (\cos \omega_0 t)^j \sin \omega t (\sin \omega_0 t)^k = c_{jk} n + O(1)$$

for some $b_{jk}, c_{jk} \in \mathbb{R}$.

Note that one immediate consequence of this is that $\forall j, k \in \mathbb{N} \cup \{0\},$ and any $\omega \in \mathbb{R}, \exists a_{jk} \in \mathbb{R}$ such that

$$\sum_{t=1}^{n} (\sin \omega t)^k (\cos \omega t)^j = a_{jk} n + O(1)$$

Corollary 1. Given that $f(\theta, t) = A \cos \omega t + B \sin \omega t$, where $A$ and $B$ take on values in a bounded set, one can conclude from Lemma 2 that for any $M \in \mathbb{N}$ and any sequence of constants $\{c_1, c_2, \ldots\} \subset \mathbb{R}$, the following limits exist:

\begin{align*}
\lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} \sum_{m=0}^{M} c_m (f(\theta_0, t))^m f(\theta, t) \\
\lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} \sum_{m=0}^{M} c_m (f(\theta, t))^m \\
\lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} \sum_{m=0}^{M} c_m (f(\theta_0, t))^m (\cos \omega_0 t)^j (\sin \omega_0 t)^k
\end{align*}

The above limits are mentioned due to the fact that they will be useful in the proof of consistency as well as the proof that the entries in the proof of (1) converges. For example, an approximation of the top left entry $d_1$ in the matrix $U$ by substituting the M-th order Taylor polynomial of $p_0 (1 - p_0)$ for $p_0 (1 - p_0)$ has a limit according to the corollary, since it has the form of (5) with $j = 2$ and $k = 0$. However, we have yet to prove that $d_1$ is well-defined, and the next corollary will accomplish that. We first observe that the above limits in Corollary 1 have the same value when we switch the two summations. Namely, for example,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} \sum_{m=0}^{M} c_m (f(\theta_0, t))^m f(\theta, t) = \sum_{m=0}^{M} c_m \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} (f(\theta_0, t))^m f(\theta, t)$$
So in fact, let’s consider instead, the right hand side of the above equality a partial sum in an infinite series indexed by $M \in \mathbb{N}$.

**Lemma 3.** *(Finite Interchangeability of Summations.)* If $X(t)$ is bounded uniformly in $t$ and $\limsup_{m \to \infty} (|a_m|)^{1/m} = 0$, then

$$
\sum_{m=0}^{\infty} a_m \frac{1}{n} \sum_{t=1}^{n} (X(t))^m = \frac{1}{n} \sum_{t=1}^{n} \sum_{m=0}^{\infty} a_m (X(t))^m
$$

**Corollary 2.** *(General Interchangeability of Summations.)* If $X(t)$ is bounded uniformly in $t$ and $\limsup_{m \to \infty} (|a_m|)^{1/m} = 0$, and the following limit is defined and finite (for instance, for when $X(t) = A \cos \omega t + B \sin \omega t$ using Lemma 2)

$$
\lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} X(t)^m
$$

then

$$
\sum_{m=0}^{\infty} a_m \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} (X(t))^m = \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} \sum_{m=0}^{\infty} a_m (X(t))^m
$$

**Corollary 3.** If for any value of $t$, $\lim_{n \to \infty} \sum_{t=1}^{n} g(t)^m h(t)$ is well-defined, and $g(t), h(t)$ are both real-valued bounded functions of $t$, and $\{a_1, a_2, \ldots\} \in \mathbb{R}$ such that $\limsup_{m \to \infty} |a_m|^{1/m} = 0$, then the following limit exists for any value of $t$:

$$
\sum_{m=0}^{\infty} a_m \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} g(t)^m h(t) = \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} \sum_{m=0}^{\infty} a_m g(t)^m h(t)
$$

The proof of existence of the left hand side proceeds by a verification of Cauchy criterion similar to that in the proof of Lemma 1. On the other hand, the equality in (9) follows from arguments similar to those in Corollary 2 and Lemma 3.

Corollary 3 basically says that we can assert the existence of the limit of $\frac{1}{n} \sum_{t=1}^{n} q(g(t)) h(t)$ wherever $g(t)$ and $h(t)$ are both bounded functions of $t$ and $q$ is a real analytic function in $g(t)$ and if the limit exists of the $\frac{1}{n} \sum_{t=1}^{n} \sum_{m=0}^{M} a_m (g(t))^m h(t)$ for all $t, m$, where $\sum_{m=0}^{M} a_m (g(t))^m$ is the $M$-th order Taylor polynomial approximation of $q(g(t))$. In particular, when viewed in conjunction with Lemma 2 and Corollary 1, we may conclude that $d_1, d_2, d_3$ are well-defined as well as all the other entries of the matrix $U$, since the function $p_{0t}(1 - p_{0t})$ is real analytic in
f(θ₀, t) for any value of θ₀. Now that the existence of the limits of the entries in the matrix 
U can be asserted, the next question is whether the entries aside from the top left four are 
correctly expressed in terms of those four (actually three: d₁, d₂, d₃ because of symmetry of 
U). To this end, note that for example,

\[ U_n(θ₀)_{33} = A^2 n^{-3} \sum_{t=1}^{n} t^2 (p₀t(1 - p₀t) \sin^2(ω₀t)) \]

\[ + B^2 n^{-3} \sum_{t=1}^{n} t^2 (p₀t(1 - p₀t) \cos^2(ω₀t)) \]

\[ - 2AB n^{-3} \sum_{t=1}^{n} t^2 (p₀t(1 - p₀t) \cos ω₀t \sin ω₀t) \]

\[ + n^{-3} \sum_{t=1}^{n} t^2 (y_t - p₀t)(A₀ \cos ω₀t + B₀ \sin ω₀t) \]

The last term above is \( O_p(n^{-1/2}) = o_p(1) \) by a use of Markov’s Inequality. Lemma 2 helps 
reduce each of the other terms. For example, using Corollary 3 and that \( \{a_m\} \) is the series 
of coefficients in the power series expansion of \( e^x \) around \( x = 0 \), we have that

\[ n^{-3} \sum_{t=1}^{n} t^2 (p₀t(1 - p₀t) \sin^2(ω₀t)) \]

\[ = n^{-3} \sum_{t=1}^{n} t^2 \sum_{m=0}^{∞} a_m (A₀ \cos ω₀t + B₀ \sin ω₀t)^m \sin^2(ω₀t) \]

\[ = \sum_{m=0}^{∞} a_m n^{-3} \sum_{t=1}^{n} t^2 (A₀ \cos ω₀t + B₀ \sin ω₀t)^m \sin^2(ω₀t) \]

\[ = \sum_{m=0}^{∞} a_m n^{-3} \sum_{t=1}^{n} t^2 (\sum_{j=1}^{J} c_{mj} \cos σ_{c,mj} t + \sum_{k=1}^{K} d_{mk} \sin σ_{d,ckt} + \sum_{l=1}^{L} b_{ml}) \]

whereas

\[ \sum_{t=1}^{n} (p₀t(1 - p₀t) \sin^2(ω₀t)) \]

\[ = \sum_{m=0}^{∞} a_m \sum_{t=1}^{n} \sum_{j=1}^{J} c_{mj} \cos σ_{c,mj} t + \sum_{k=1}^{K} d_{mk} \sin σ_{d,ckt} + \sum_{l=1}^{L} b_{ml} \]

where the representations in the last expression arise from the fact that for each \( m \), the 
quantity \( (A₀ \cos ω₀t + B₀ \sin ω₀t)^m \sin^2(ω₀t) \) according to the proof of Lemma 2 can be reduced
to a linear combination of finitely many terms in the form (1) of a constant, (2) \( \cos(\sigma t) \),
or (3) \( \sin(\sigma t) \) for some value \( \sigma \in \mathbb{R} \). Note that the terms in the linear combination do
not necessarily have the same value of \( \sigma \). On the other hand, \( n^{-3}\sum_{t=1}^{n} t^2 \cos(\sigma t) \to 0 \),
and \( n^{-3}\sum_{t=1}^{n} t^2 \sin(\sigma t) \to 0 \) for any \( \sigma \neq 0(\text{mod} \ 2\pi) \), whereas if \( \sigma = 0(\text{mod} \ 2\pi) \),
then
\[
\lim_{n \to \infty} n^{-3}\sum_{t=1}^{n} t^2 \cos(\sigma t) = \frac{1}{3} d_2
\]
Similarly,
\[
\lim_{n \to \infty} n^{-3}\sum_{t=1}^{n} t^2 p_{0t}(1 - p_{0t}) \sin^2(\omega_0 t) = \frac{1}{3} d_3
\]
Performing a similar argument on the other entries of the matrix \( U_n(\theta_0) \) results in the
representation of the matrix \( U \) in (1).

**Lemma 4.** The matrix \( U \)'s determinant is
\[
\frac{1}{12}(d_1 d_2 - d_3^2)(A_0^2 d_2 + B_0^2 d_1 - 2A_0 B_0 d_3)
\]
In addition, \( U \) is positive-definite.

2.2. Consistency

Let \( \hat{\theta}_n \) be a minimizer of \( G_n(\theta) \). Note that if we can show that \( \hat{\theta}_n^* := \text{diag}(1, 1, n)(\hat{\theta}_n - \theta_0) = o_p(1) \), then it follows that \( \hat{\theta}_n = o_p(1) \). We will prove the former, stronger consistency.

**Lemma 5 (Sufficient Condition for Weak Consistency of \( \hat{\theta}_n^* \)).** Suppose that for any
fixed \( \delta > 0 \), the following is true for some \( W_n(\delta) \) converging to 0 in probability uniformly in \( \theta^* \), and \( K_n(\delta) > K(\delta) \) for some \( K(\delta) > 0 \) uniformly in \( n \),

\[
\inf_{||\theta^*|| \geq \delta} (G_n^*(\theta^*) - G_n^*(0)) = W_n(\delta) + K_n(\delta)
\]
then
\[
\hat{\theta}_n^* = o_p(1).
\]
where

\[ \theta^* := \text{diag}(1, 1, n)(\theta - \theta_0) \]

\[ G_n^*(\theta^*) := G_n(\text{diag}(1, 1, 1/n)\theta^* + \theta_0) = G_n(\theta) \]

Note that also \( G_n^*(0) = G_n(\theta_0) \), and it follows that \( \hat{\theta}_n \) minimizes \( G_n(\theta) \) if and only if \( \hat{\theta}_n^* \) minimizes \( G_n^*(\theta^*) \).

\[
\frac{1}{n} G_n^*(\theta^*) - \frac{1}{n} G_n^*(0) = \frac{1}{n} \{ G_n(\theta) - G_n(\theta_0) \} \\
= -\frac{1}{n} \sum_{t=1}^{n} (y_t - p_{0t})(f(\theta, t) - f(\theta_0, t)) - \frac{1}{n} \sum_{t=1}^{n} p_{0t}(f(\theta, t) - f(\theta_0, t)) \\
+ \frac{1}{n} \sum_{t=1}^{n} \log(1 + e^{f(\theta, t)}) - \frac{1}{n} \sum_{t=1}^{n} \log(1 + e^{f(\theta_0, t)})
\]

Taylor-expanding \( \frac{1}{n} \sum_{t=1}^{n} \log(1 + e^{f(\theta, t)}) \) around \( u(t) = f(\theta_0, t) \) for each \( t \) gives

\[
\frac{1}{n} \sum_{t=1}^{n} \log(1 + e^{f(\theta, t)}) = \\
\frac{1}{n} \sum_{t=1}^{n} \log(1 + e^{f(\theta_0, t)}) + \frac{1}{n} \sum_{t=1}^{n} p_{0t}(f(\theta, t) - f(\theta_0, t)) + \frac{1}{n} \sum_{t=1}^{n} \tilde{p}_t(1 - \tilde{p}_t)(f(\theta, t) - f(\theta_0, t))^2
\]

with each \( \tilde{p}_t := \frac{e^{f(\theta_0, t)}}{1 + e^{f(\theta_0, t)}} \), where \( \tilde{\theta}_t \) is a value lying between \( \theta_0 \) and \( \theta \), varies according as \( t \) and exists according to the mean value theory. Thus

\[
\frac{1}{n} \{ G_n(\theta) - G_n(\theta_0) \} = -\frac{1}{n} \sum_{t=1}^{n} (y_t - p_{0t})(f(\theta, t) - f(\theta_0, t)) + \frac{1}{n} \sum_{t=1}^{n} \tilde{p}_t(1 - \tilde{p}_t)(f(\theta, t) - f(\theta_0, t))^2
\]

Denote \( \sum_{t=1}^{n} \tilde{p}_t(1 - \tilde{p}_t)(f(\theta, t) - f(\theta_0, t))^2 \) from the above display by \( R_n(\theta, \theta_0) \).

**Lemma 6.** To show that for any \( \delta > 0, (10) \) holds, it suffices to show that

\[
\lim_{n \to \infty} \inf_{\|\theta^*\| \geq \delta} n^{-1} R_n(\theta, \theta_0) > 0 \tag{11}
\]

**Remark 1.** In the proof of Lemma (6), it is shown that \( -\frac{1}{n} \sum_{t=1}^{n} (y_t - p_{0t})(f(\theta, t) - f(\theta_0, t)) = o_p(1) \) uniformly in \( \theta \), where the amplitude components of \( \theta^* = \text{diag}(1, 1, n)(\theta - \theta_0) \) are contained in a compact subset of \( \mathbb{R} \). The uniformity is important in satisfying the representation
of \( \inf_{\|\theta^*\| \geq \delta} (G_n^*(\theta^*) - G_n^*(0)) \) as a sum of \( W_n(\delta) \) and \( K_n(\delta) \), with \( W_n(\delta) \) converging to 0 in probability uniformly in \( \theta^* \). In particular, note that

\[
\inf_{\|\theta^*\| \geq \delta} (G_n^*(\theta^*) - G_n^*(0)) = \inf_{\|\theta^*\| \geq \delta} \left( -\frac{1}{n} \sum_{t=1}^{n} (y_t - p_{0t})(f(\theta, t) - f(\theta_0, t)) + \frac{1}{n} \sum_{t=1}^{n} \tilde{p}_t(1 - \tilde{p}_t)(f(\theta, t) - f(\theta_0, t))^2 \right)
\]

and

\[
\sup_{\|\theta^*\| \geq \delta} \left( -\frac{1}{n} \sum_{t=1}^{n} (y_t - p_{0t})(f(\theta, t) - f(\theta_0, t)) + \frac{1}{n} \sum_{t=1}^{n} \tilde{p}_t(1 - \tilde{p}_t)(f(\theta, t) - f(\theta_0, t))^2 \right) \geq \inf_{\|\theta^*\| \geq \delta} \left( -\frac{1}{n} \sum_{t=1}^{n} (y_t - p_{0t})(f(\theta, t) - f(\theta_0, t)) + \frac{1}{n} \sum_{t=1}^{n} \tilde{p}_t(1 - \tilde{p}_t)(f(\theta, t) - f(\theta_0, t))^2 \right)
\]

The above chain of inequalities following from the fact that \( \inf(a(x) + b(x)) \geq \inf a(x) + \inf b(x) \) for all real-valued functions \( a, b \) defined on the same domain. To finish proving that the representation of \( \inf_{\|\theta^*\| \geq \delta} (G_n^*(\theta^*) - G_n^*(0)) \) as in (10) is attained, it remains to show (11) for any fixed \( \delta > 0 \).

**Proposition 1 (Consistency).** For \( \Theta := \{ \theta \in \Theta_0 : A, B \in \mathcal{K}, \omega \in \Omega \} \), where \( \mathcal{K} \) is a compact set in \( \mathbb{R} \), \( \Omega := \{ \omega : c \leq \omega \leq \pi - c \} \) for some \( 0 < c < \pi/2 \), the display (11) holds in probability for any fixed \( \delta > 0 \), which when combined with the result of Lemma 6 shows that \( \hat{\theta}_n^* = o_p(1) \).

**Proof.** To establish (11), note that \( \tilde{p}_t(1 - \tilde{p}_t) \) is bounded below by a \( d := \frac{\varepsilon^K}{(1 + \varepsilon^K)^2} > 0 \), where \( K < \infty \) is a fixed number such that \( \sup_{\theta}(|A| + |B|) \leq K \), such a \( d \) exists when \( A, B \) take on values inside a compact subset of \( \mathbb{R} \). This is because \( p_t = 1 - \frac{1}{e^{\lambda t\cos\omega t} + e^{\lambda t\sin\omega t}} \) is bounded away from 0 and 1 when \( A, B \) are bounded in absolute value. Therefore it suffices to show that

\[
\liminf_{n \to \infty} \inf_{\|\theta^*\| \geq \delta} \frac{1}{n} \sum_{t=1}^{n} (f(\theta, t) - f(\theta_0, t))^2 > 0
\]

Walker’s approach [11] in obtaining the consistency of the minimizer of the residual sum of squares for the harmonic regression problem \( y_t = A_0 \cos(\omega_0 t) + B_0 \sin(\omega_0 t) + \epsilon_t \) for i.i.d.
noise $\epsilon_t$ centers on first maximizing the periodogram as a function of $\omega$. Song and Li [8] made mathematically explicit the relation between the regression sum of squares $SSR(\omega)$ as asymptotically equal to the periodogram as a function of $\omega$, uniformly in $\omega$. Their respective techniques will be applied in the proof of (11). Before proceeding, note the following:

$$\lim_{n \to \infty} \inf_{||\theta^*|| \geq \delta} \frac{1}{n} \sum_{t=1}^{n} (f(\theta, t) - f(\theta_0, t))^2$$

$$= \lim_{n \to \infty} \inf_{|n(\omega-\omega_0)| \geq \delta_0} \frac{1}{n} \sum_{t=1}^{n} (f(\theta, t) - f(\theta_0, t))^2$$

$$\geq \lim_{n \to \infty} \inf_{|n(\omega-\omega_0)| \leq \delta_0} \frac{1}{n} \sum_{t=1}^{n} (f(\theta, t) - f(\theta_0, t))^2$$

The following will be pursued separately, for an appropriately(to be decided in the proof) small $\delta_0 > 0$:

(13) \[ \lim_{n \to \infty} \inf_{|n(\omega-\omega_0)| \geq \delta_0} \frac{1}{n} \sum_{t=1}^{n} (f(\theta, t) - f(\theta_0, t))^2 > 0 \]

(14) \[ \lim_{n \to \infty} \inf_{|n(\omega-\omega_0)| \leq \delta_0} \frac{1}{n} \sum_{t=1}^{n} (f(\theta, t) - f(\theta_0, t))^2 > 0 \]

**Proof.** To prove (13), note that for each fixed $\omega$, the minimum of $\sum_{t=1}^{n} (f(\theta, t) - f(\theta_0, t))^2$ occurs at $\tilde{\beta} := \tilde{\beta}(\omega) := (X'X)^{-1}X'X_0\beta_0$ by a calculus argument, where $X := X(\omega) = (X_1(\omega), \ldots, X_n(\omega))'$, with $X_t(\omega) := (\cos(\omega t), \sin(\omega t))'$, and $X_0 := X(\omega_0), \beta_0 := (A_0, B_0)'$. Denote by $\theta(\omega)$ the set of all $\theta$ whose frequency component is $\omega$. Then for fixed $\omega$,

$$\min_{\theta \in \theta(\omega)} \sum_{t=1}^{n} (f(\theta, t) - f(\theta_0, t))^2 = ||X_0\beta_0||^2 - ||X\tilde{\beta}||^2$$

Hence

$$\lim_{n \to \infty} \inf_{|n(\omega-\omega_0)| \geq \delta_0} \frac{1}{n} \sum_{t=1}^{n} (f(\theta, t) - f(\theta_0, t))^2$$

$$\geq \lim_{n \to \infty} \inf_{|n(\omega-\omega_0)| \geq \delta_0} \left( n^{-1} ||X_0\beta_0||^2 - n^{-1} ||X\tilde{\beta}||^2 \right)$$
\[
\lim_{n \to \infty} \frac{1}{n} \left| \begin{array}{c}
X_0 \beta_0 \end{array} \right|^2 - \limsup_{n \to \infty} \sup_{|n(\omega - \omega_0)| \geq \delta_0} \left( n^{-1} \left| X_{\beta} \right|^2 \right) = A_0^2 + B_0^2 - \limsup_{n \to \infty} \sup_{|n(\omega - \omega_0)| \geq \delta_0} \frac{2}{n^2} \left| X'X_0 \beta_0 \right|^2 (1 + e_n)
\]

with the last equality following from Song & Li's [8] argument by linear algebra, where 
\[e_n = \frac{1}{2} \left( X_0 \beta_0 \right)'XR_nX'(X_0 \beta_0)/\left| X'X_0 \beta_0 \right|^2,\]
and the symmetric matrix \( R_n = o(1) \) uniformly in \( \omega \in \Omega \), which is a compact set bounded away from 0 and \( \pi \), and from which \( e_n = o(1) \) uniformly in \( \omega \in \Omega \) follows. In addition, if one defines the following as in Walker [11]:

\[
\begin{align*}
D_0 &:= \frac{1}{2} (A_0 - iB_0) \quad D_0^* := \frac{1}{2} (A_0 + iB_0) \\
M_n(u) &:= \sum_{t=1}^{n} e^{iut} = \begin{cases} 
\frac{e^{1/2-i(n+1)u} \sin(1/2nu)}{\sin(u/2)}, & u \in (0, 2\pi) \\
n & u = 0 \text{ or } 2\pi 
\end{cases}
\end{align*}
\]

Then using the fact pointed out by Walker [11] that
\[
\max_{0 \leq \omega \leq \pi} \left| M_n(\omega + \omega_0) \right| = O(1), \quad \max_{0 \leq \omega \leq \pi} \left| M_n(\omega - \omega_0) \right| = n
\]

One has
\[
\frac{1}{n^2} \left| X'X_0 \beta_0 \right|^2 (1 + e_n) = \frac{1}{n^2} \left| D_0 M_n(\omega + \omega_0) + D_0^* M_n(\omega - \omega_0) \right|^2 (1 + e_n) = \frac{1}{n^2} \left( O(1) + O(n) + \left| D_0^* M_n(\omega - \omega_0)^2 \right| \right) (1 + e_n) = \left( \frac{A_0^2 + B_0^2}{4} \right) \left| M_n(\omega - \omega_0) \right|^2 + o(1) (1 + e_n)
\]

For all \( \eta > 0 \) small enough, \( (\sin(\eta/2)/(\eta/2))^2 > 1/(\pi^2) \). Let \( \delta_0 \leq 1 \) be one such value of \( \eta \). Then Walker [11] noted that the following is true for all \( n \) sufficiently large,

\[
\frac{1}{n^2} \max_{|n(\omega - \omega_0)| \geq \delta_0} \left| M_n(\omega - \omega_0) \right|^2
\]
\[
\frac{1}{n^2} \left( \sin(\delta_0/2)/\sin(n^{-1}\delta_0/2) \right) \\
= \sin^2(\delta_0/2)/(n^2 \sin^2(n^{-1}\delta_0/2)) \\
\rightarrow \sin^2(\delta_0/2) \cdot \lim_{n \to \infty} \frac{1}{n^2 \sin^2(n^{-1}\delta_0/2)} \\
< \frac{\delta_0^2}{4} \cdot \frac{1}{\delta_0^2/4} = 1
\]

So then \( \limsup_{n \to \infty} \sup_{|n(\omega - \omega_0)| \geq \delta_0} \left( n^{-1} \left\| X\tilde{\beta} \right\|^2 \right) = \limsup_{n \to \infty} \sup_{|n(\omega - \omega_0)| \geq \delta_0} \frac{2}{n^2} \left\| X'X_0\beta_0 \right\|^2 < \frac{A_0^2 + B_0^2}{2} \) and (13) follows.

Note that in the preceding computations, a \( \delta_0 > 0 \) suitably small was chosen. To prove (14) for all \( \delta_0 \) small enough, care will be taken at the end of the following proof to select a \( \delta_0 \) which depends on \( \delta \) subject to the constraint of not exceeding the previous \( \delta_0 \). Without loss of generality assume \( \delta_0 \leq \min \{ 1, \delta/2 \} \).

Then
\[
\frac{1}{n} \sum_{t=1}^{n} (f(\theta, t) - f(\theta_0, t))^2 \\
= \frac{1}{n} \sum_{t=1}^{n} (A \cos \omega t - A_0 \cos \omega_0 t)^2 + \frac{1}{n} \sum_{t=1}^{n} (B \sin \omega t - B_0 \sin \omega_0 t)^2 \\
+ AB \frac{1}{n} \sum_{t=1}^{n} 2 \sin \omega t \cos \omega_0 t + A_0 B_0 \frac{1}{n} \sum_{t=1}^{n} 2 \sin \omega_0 t \cos \omega_0 t \\
- AB_0 \frac{1}{n} \sum_{t=1}^{n} 2 \cos \omega t \sin \omega_0 t + A_0 B \frac{1}{n} \sum_{t=1}^{n} 2 \cos \omega_0 t \sin \omega_0 t
\]

Let \( K > 0 \) be an upper bound for all possible values of \( |A| \) and \( |B| \). \( K \) exists because \( A, B \) take on values in a compact set. Let \( \omega \) satisfy the additional requirement that \( |n(\omega - \omega_0)| \leq \delta_0 \) for some \( \delta_0 > 0 \). Then note that
\[
\frac{1}{n} \sum_{t=1}^{n} \cos 2\omega t = o(1) \quad \frac{1}{n} \sum_{t=1}^{n} \sin 2\omega t = o(1) \\
\frac{1}{n} \sum_{t=1}^{n} \cos(\omega + \omega_0) t = o(1) \quad \frac{1}{n} \sum_{t=1}^{n} \sin(\omega + \omega_0) t = o(1)
\]

uniformly in \( \omega \in \Omega \). Note also that
\[
2 \sin \omega t \cos \omega_0 t = \sin(\omega + \omega_0) t + \sin(\omega - \omega_0) t
\]
\[ 2\sin \omega_0 \cos \omega t = \sin(\omega + \omega_0)t - \sin(\omega - \omega_0)t \]

Moreover, by the mean value theory, for all \( t \), \( \exists \omega_t \) between \( \omega_0 \) and \( \omega \) such that

\[
\frac{1}{n} \sum_{t=1}^{n} (1 - \cos(\omega - \omega_0)t) = \frac{1}{n} \sum_{t=1}^{n} t(\sin(\omega_t - \omega_0)t)(\omega - \omega_0)
\]

Since \( |n(\omega - \omega_0)| \leq \delta_0 \leq 1 \), we have \( |\sin(\omega - \omega_0)t| < |\omega - \omega_0|t \), and the absolute value of the above does not exceed \( n^2(\omega - \omega_0)^2/2 \leq \delta_0^2/2 \) for all \( n \) large enough. On the other hand,

\[
\left| \frac{1}{n} \sum_{t=1}^{n} \sin(\omega - \omega_0)t \right| \leq \frac{1}{n} \sum_{t=1}^{n} |n(\omega - \omega_0)| \leq \delta_0
\]

Therefore, using the assumption that \( A \) takes on values in a compact set in \( \mathbb{R} \), the following is true uniformly in \( \omega \):

\[
\frac{1}{n} \sum_{t=1}^{n} (A\cos \omega t - A_0\cos \omega_0 t)^2
\]

\[
= \frac{1}{n} \sum_{t=1}^{n} A^2(\cos \omega t)^2 + \frac{1}{n} \sum_{t=1}^{n} A_0^2(\cos \omega_0 t)^2 - AA_0 \frac{1}{n} \sum_{t=1}^{n} 2\cos \omega t \cos \omega_0 t
\]

\[
= A^2(1/2 + o(1)) + A_0^2(1/2 + o(1)) - AA_0
\]

\[
+ AA_0[(1 - \frac{1}{n} \sum_{t=1}^{n} \cos(\omega - \omega_0)t) - \frac{1}{n} \sum_{t=1}^{n} \cos(\omega + \omega_0)t]
\]

\[
= \frac{1}{2} (A - A_0)^2 + AA_0[1 - \frac{1}{n} \sum_{t=1}^{n} \cos(\omega - \omega_0)t] + o(1)
\]

\[
\geq \frac{1}{2} (A - A_0)^2 - |AA_0| \cdot \frac{\delta_0^2}{2} + o(1)
\]

\[
\geq \frac{1}{2} (A - A_0)^2 - K^2 \cdot \frac{\delta_0^2}{2} + o(1)
\]

for all \( n \) large enough. Similarly, uniformly in \( \omega \), for all \( n \) large enough, the following holds:

\[
\frac{1}{n} \sum_{t=1}^{n} (B\sin \omega t - B_0\sin \omega_0 t)^2 \geq \frac{1}{2} (B - B_0)^2 - K^2 \cdot \frac{\delta_0^2}{2} + o(1)
\]

Moreover,

\[
\lim_{n \to \infty} \left| AB_0 \frac{1}{n} \sum_{t=1}^{n} 2\cos \omega t \sin \omega_0 t \right| \leq K^2 \lim_{n \to \infty} \left| \frac{1}{n} \sum_{t=1}^{n} \sin(\omega_0 + \omega)t + \frac{1}{n} \sum_{t=1}^{n} \sin(\omega - \omega_0)t \right| \leq K^2 \delta_0
\]
Similarly,

\[ \lim_{n \to \infty} \left| A_0 B \frac{1}{n} \sum_{t=1}^{n} 2\cos \omega_0 t \sin \omega t \right| \leq K^2 \delta_0 \]

So

\[ \frac{1}{n} \sum_{t=1}^{n} (f(\theta, t) - f(\theta_0, t))^2 \geq \frac{1}{2} (A - A_0)^2 + \frac{1}{2} (B - B_0)^2 - K^2 \delta_0^2 - 2K^2 \delta_0 + o(1) \]

But

\[ ||\theta^*|| \geq \delta \text{ implies } (A - A_0)^2 + (B - B_0)^2 + n^2(\omega - \omega_0)^2 \geq \delta^2 \]

Knowing that \( n(\omega - \omega_0) \leq \delta_0 \) gives

\[ \frac{1}{2} (A - A_0)^2 + \frac{1}{2} (B - B_0)^2 \geq \frac{\delta^2 - \delta_0^2}{2} \]

In order that (14) holds, it suffices to have

\[ \frac{\delta^2 - \delta_0^2}{2} > K^2 \delta_0^2 + 2K^2 \delta_0 \]

which is true if

\[ \frac{\delta^2 - \delta_0}{2} > 3K^2 \delta_0 \]

since we have assumed without loss of generality that \( \delta_0 \leq 1 \). Solving the above inequality gives

\[ \delta_0 < \frac{\delta^2}{6K^2 + 1} \]

So \( \delta_0 \) can be chosen to be \( \min \left\{ \frac{\delta^2}{6K^2 + 2}, 1, \delta/2 \right\} \). □

2.3. Rates of Convergence

We will specifically identify the rates of convergence of the components of the parameter. Recall that

\[ U_n(\theta_0) \to_p U \]

Let \( U_n(\theta_0) = U^*_n(\theta_0) + V_n(\theta_0) \), where \( U^*_n(\theta_0) \) is the matrix whose entries consist of deterministic terms in the corresponding entries in \( U_n(\theta_0) \) and therefore \( V_n(\theta_0) \) consist only of the stochastic components of the entries. We would like to prove that:
Lemma 7. Assuming that $\hat{\theta}_n^* = o_p(1)$ and that all possible values of $A, B$ are attained in a compact set, say $\mathcal{K}$, then the following is true uniformly in $\tilde{\theta}_n := \theta_0 + s(\hat{\theta}_n - \theta_0)$ for all $s \in [0, 1]$

$$U_n(\hat{\theta}_n) - U_n(\theta_0) \rightarrow P 0$$

where $U_n(\hat{\theta}_n)$ is symmetric and

$$U_n(\hat{\theta}_n)_{11} = n^{-1} \sum_{t=1}^{n} \tilde{p}_t (1 - \tilde{p}_t)(\cos \tilde{\omega}_n t)^2$$

$$U_n(\hat{\theta}_n)_{12} = n^{-1} \sum_{t=1}^{n} \tilde{p}_t (1 - \tilde{p}_t)\cos \tilde{\omega}_n t \sin \tilde{\omega}_n t$$

$$U_n(\hat{\theta}_n)_{13} = n^{-2} \sum_{t=1}^{n} \tilde{p}_t (1 - \tilde{p}_t)(-A \sin \tilde{\omega}_n t + B \cos \tilde{\omega}_n t)\cos \tilde{\omega}_n t + t(y_t - \tilde{p}_t) \sin \tilde{\omega}_n t$$

$$U_n(\hat{\theta}_n)_{22} = n^{-1} \sum_{t=1}^{n} \tilde{p}_t (1 - \tilde{p}_t)(\sin \tilde{\omega}_n t)^2$$

$$U_n(\hat{\theta}_n)_{23} = n^{-2} \sum_{t=1}^{n} \tilde{p}_t (1 - \tilde{p}_t)(-A \sin \tilde{\omega}_n t + B \cos \tilde{\omega}_n t)\sin \tilde{\omega}_n t - t(y_t - \tilde{p}_t) \cos \tilde{\omega}_n t$$

$$U_n(\hat{\theta}_n)_{33} = n^{-3} \sum_{t=1}^{n} \tilde{p}_t (1 - \tilde{p}_t)(-A \sin \tilde{\omega}_n t + B \cos \tilde{\omega}_n t)^2 t^2 + t^2(y_t - \tilde{p}_t)(A \cos \tilde{\omega}_n t + B \sin \tilde{\omega}_n t)$$

Proposition 2 (Rates of Convergence). The joint rate of convergence of the parameter $\hat{\theta}_n^*$ is given by: $\hat{\theta}_n^* = O_p(n^{-1/2})$.

Now that the consistency of the estimators $\hat{\theta}_n$ of $\theta_0$ has been established, and the second order derivative matrix has been shown to converge in probability to a positive definite symmetric matrix based on the consistency of the estimator $\hat{\theta}_n$ satisfying $(\hat{A}_n - A_0, \hat{B}_n - B_0, n(\hat{\omega}_n - \omega_0))' = \hat{\theta}_n^* = o_p(1)$, and using that $U_n(\theta) = D_n \nabla^2 G_n(\theta)D_n$, we obtain that the centered criterion

$$0 \geq \frac{1}{n}(G_n(\hat{\theta}_n) - G_n(\theta_0)) = \nabla G_n(\theta_0)'(\theta - \theta_0) + (\theta - \theta_0)'H_n(\theta)(\theta - \theta_0)$$

$$= n^{-1/2}(\hat{\theta}_n^*)'D_n \nabla G_n(\theta_0) + (\hat{\theta}_n^*)'D_n H_n(\hat{\theta}_n)D_n(\hat{\theta}_n^*)$$
Define the notation \( \gtrsim \) as in Radchenko [7] with \( a(\theta) \gtrsim b(\theta) \) meaning there exists a positive \( c_0 \) such that \( a(\theta) \geq c_0 b(\theta) \) for all \( \theta \) in a sufficiently small neighborhood of the origin. Recall that we defined \( U_n(\theta) := D_n \nabla^2 G_n(\theta) D_n \), and proved previously that \( U_n(\hat{\theta}_n) \to_p U \). So using the symbolic representations as in Radchenko[7], we can define the following

\[
\mathcal{M}_n(\hat{\theta}_n^*) := (\hat{\theta}_n^*)' D_n H_n(\hat{\theta}_n) D_n(\hat{\theta}_n^*) = (\hat{\theta}_n^*)' D_n \int_0^1 (1 - s) \nabla^2 G_n(\theta_0 + s(\hat{\theta}_n - \theta_0)) \, ds \, D_n(\hat{\theta}_n^*)
\]

\[
\mathcal{N}_n(\hat{\theta}_n^*) := -n^{-\frac{3}{2}} (\hat{\theta}_n^*)' D_n \nabla G_n(\theta_0) = -n^{-1/2} (\hat{\theta}_n^*)' (Z_{n1}, Z_{n2}, Z_{n3})
\]

Then

\[
0 \geq \frac{1}{n} (G_n(\hat{\theta}_n) - G_n(\theta_0)) = \mathcal{M}_n(\hat{\theta}_n^*) - \mathcal{N}_n(\hat{\theta}_n^*)
\]

\[
\mathcal{M}_n(\hat{\theta}_n^*) \to_p \frac{1}{2} \hat{\theta}_n^* U \hat{\theta}_n^* \gtrsim \| \hat{\theta}_n^* \|^2 \quad \text{with probability tending to 1.}
\]

The convergence of \( \mathcal{M}_n(\hat{\theta}_n^*) \) (with inner probability tending to 1) is due to (2) (uniform convergence in \( s \) in probability) and the \( \gtrsim \) due to \( U \) being positive definite. Also, \( [\mathcal{N}_n(\hat{\theta}_n^*)]^{+} = O_p(n^{-1/2}) \| \hat{\theta}_n^* \| \) using either Markov’s inequality or the uniform tightness of the vector \( D_n \nabla G_n(\theta_0) \), which will be proved later as a consequence of its weak convergence to a random vector in \( \mathbb{R}^3 \). Thus the conditions in Radchenko’s Lemma 1 for obtaining the rates of convergence are satisfied. In particular, \( \| \hat{\theta}_n^* \| = O_p(n^{-1/2}) \). From which we conclude

\[
(\hat{A}_n - A_0, \hat{B}_n - B_0, n(\hat{\omega}_n - \omega_0))' = O_p(n^{-1/2})
\]
CHAPTER 3

MONTE CARLO SIMULATION AND RESULTS

The goal of the Monte Carlo simulations is to see how closely the simulations agree with theoretical derivations and to gauge likely size of error for various sample sizes, if the model is a good fit for the data.

3.1. The Monte Carlo simulation of time series

Using an underlying parameter \([2;4;0.4]\), Monte Carlo simulations were performed in MATLAB [5] to generate 10 files each consisting of 1000 identical-length vectors, whose lengths are reflected in the filenames, ranging from length \(n = 100\) to \(n = 1000\).

Each vector consists of \(n\) entries of 1’s and 0’s, determined by comparing the outcome of a random uniform variable on \([0,1]\) with the Bernoulli random variable parameter

\[ p_t = \frac{\exp(2 \cos(.4t) + 4 \sin(.4t))}{1 + \exp(2 \cos(.4t) + 4 \sin(.4t))}. \]

If the random uniform variable assumes a value < \(p_t\), then the corresponding entry is 1, otherwise it is 0.

3.2. Using the MultiStart framework

Next, we begin a random search over a compact region containing the true parameter \([2;4;0.4]\) for a minimum within the region.

The implementation is carried out with the aid of MATLAB’s GLocal Optimization Toolbox. To test its correctness and accuracy, several simulations with only one amplitude component in the model were performed. The choice of only one amplitude component allows visualizations of the location of the minimizer. In all these preliminary simulations and with 600 uniformly generated initial start points, the location of the minimum is found to a high degree of precision.

MATLAB’s implementation of the MultiStart algorithm is based on Ugray et.al [10]’s discussion of the MultiStart algorithm, which consists of two phases - the global search phase, and the local search phase. This idea for random MultiStart and two-phase search was called Multilevel Single Linkage (MLSL) and discussed thoroughly in Kan & Timmer
[3] and [4]. Kan & Timmer points out that even though for continuous criterion functions, the process of using uniformly generated start points will produce the global minimum with probability converging to 1 as the number of starting points goes to infinity, it is very inefficient. In addition, they note that the so-called Grid Search, in which the function is evaluated in each point of a regular grid, fares even worse according to several probabilistic criteria based on analyses by various authors. Kan & Timmer’s method and its extensions are discussed in [10], whose main results suggest that the MultiStart framework results in a rather efficient search for the global minimum.

In our search for the global minimum, at least 600 initial random start points are generated. Specifically, the active-set algorithm is used with the fmincon solver for the locations of the minima. Each call to the MultiStart procedure results in an output of a list of local minima found, ordered by ascending criterion function values.

3.3. Refining the search globally and locally using gradient and Hessian information

1. The coarse run. The results from a preliminary search for minimizers of the objective on shorter time series sized 50 suggests augmenting the search region for the amplitude components since the amplitude components of the candidate global minimizers frequently hit the boundary when the original search intervals was [-10,10] for the amplitude components and [0.01, 3.15] for the frequency components. Moreover, many more outliers occur on the positive side of zero for the estimates of the amplitude components. On the other hand, the estimate of the frequency component stays no farther than .5 from 0.4. This led to a change of the initial interval to be [-4,20] and [0, 40], and [pts(3)-.1,pts(3)+.1], where pts is the candidate argmin obtained.
in the previous run, especially if the value of the function is less than .12. Note that this threshold was obtained by examining the case where one of the parameter components hits the boundary, in which case the function value is below .1. On the other hand, this threshold .12 would be expected to vary depending on the data and the underlying parameter, here [2;4;.4] in particular). This is reflected in the file NRefined100.m and NRefined200.m. Simulations suggest that beginning with sample sizes 300 and above, the estimates of the parameters fall in the range [0,5] and [2,7] respectively, so instead the search intervals for time series longer than 300 is specified to be [pts(1)-.5,pts(1)+.5], [pts(2)-.5,pts(2)+.5], and [pts(3)-.1, pts(3)+.1]. Note there that the search intervals for the amplitude components are much smaller, so that a global argmin can be more effectively sought. This is reflected in NRefined300.m through NRefined1000.m.

3. The third run: The minimizers from the above are saved in text files called NRefined%drun1000nst600r.txt. They consist of vector rows each used as an initial point in the search for global minimum based on the corresponding y-vector, over the interval [pts(1)-5, pts(1)+5], [pts(2)-5, pts(2)+5], and [pts(3)-1, pts(3)+1]. For the time series of length 400 and greater, the intervals [pts(1)-2, pts(1)+2], [pts(2)-2, pts(2)+2], and [pts(3)-.1, pts(3)+.1] are used instead. This will produce dozens of local argmins as determined by the Matlab’s MultiStart procedure. Of these, the best ten, as measured by the smallest objective values, are chosen. Using each of these ten points as initial points x, the MATLAB searches for local minimum within the interval [x(1)-.01,x(1)+.01], [x(2)-.01,x(2)+.01], and [x(3)-.0001,x(3)+.0001]. The results of these searches are saved in 3 files for each time series. The first is the best guesses of argmins, and the second the norm of the gradient at the best guess, and the third the best 10 guesses of argmins for each vector y.

4. The fourth run: This uses a concentrated interval [pts(1)-.01, pts(1)+.01], [pts(2)-.01, pts(2)+.01], [pts(3)-.0001,pts(3)+.0001], and the same number of starting points, along with the last output file above as initial starting points, which is guaranteed to produce no worse results than the previous run, if not better.

5. Collecting simulation results: Three sets of files are again obtained as in the
previous step, first consisting of the best guesses of the argmins, the second the norms of the gradients at the best guesses, and the third the best ten guesses of argmins for each vector y. These are saved in case further runs based on the same set of vectors y are deemed profitable.

3.4. Tools, technical notes, and data analysis

The programs running the simulations and searches for the global minimizers were written in MATLAB and carried out with the aid of its Global Optimization Toolbox, and the outputs are graphed and analyzed using the R language [6]. Moreover, the simulations were run on UNT’s High Performance Computing clusters Talon 1.0. Although more data could have been obtained and analyzed, it seems that the accuracy of the searches for the minimizers could be optimized by looking at the output at each stage of the refinement discussed above, and there often seems to be room for improvement first and second steps above. In particular, this entails multiple experiments checking for a heuristically advantageous choice of search intervals, examining how closely the top ten or twenty minimizers found in each run agree with each other, evaluating how increasing the size of initial start point sets and changing the number of best minimizers included in one run as starting points contribute to the improvement of the estimate of the argmin in the next run.

Data Analysis Result: Improvement of the RMSE:
The root mean square error (RMSE) steadily improves among the parameter component estimates except for a slight anomaly: the RMSE for the second amplitude parameter is smaller at n=800 than at n=900.

On the next few pages, the RMSE, bias, and variance for each component estimate of the parameter are plotted against the lengths of time series in the figures, and the corresponding values and univariate normality test results are displayed in the tables.
Figure 3.1. Root mean square error vs. Length of time series
Bias of Estimates vs. Time Series Length

Figure 3.2. Bias vs. Length of time series
Figure 3.3. Variance of estimates vs. Length of time series
### Table 3.1. Variance vs. Time series length

<table>
<thead>
<tr>
<th>n</th>
<th>VAR($\hat{A}_{n,m}$)</th>
<th>VAR($\hat{B}_{n,m}$)</th>
<th>VAR($\hat{\omega}_{n,m}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>1.453547000</td>
<td>1.955219000</td>
<td>0.000001042</td>
</tr>
<tr>
<td>200</td>
<td>0.390238100</td>
<td>0.458756020</td>
<td>0.000001348</td>
</tr>
<tr>
<td>300</td>
<td>0.251101500</td>
<td>0.277695800</td>
<td>0.000000365</td>
</tr>
<tr>
<td>400</td>
<td>0.170453100</td>
<td>0.198571800</td>
<td>0.000000148</td>
</tr>
<tr>
<td>500</td>
<td>0.133900700</td>
<td>0.147118500</td>
<td>0.000000074</td>
</tr>
<tr>
<td>600</td>
<td>0.106458500</td>
<td>0.121864600</td>
<td>0.000000039</td>
</tr>
<tr>
<td>700</td>
<td>0.098857130</td>
<td>0.102333800</td>
<td>0.000000028</td>
</tr>
<tr>
<td>800</td>
<td>0.089150110</td>
<td>0.083326700</td>
<td>0.000000019</td>
</tr>
<tr>
<td>900</td>
<td>0.074343520</td>
<td>0.085416640</td>
<td>0.000000013</td>
</tr>
<tr>
<td>1000</td>
<td>0.060243710</td>
<td>0.070195660</td>
<td>0.000000009</td>
</tr>
</tbody>
</table>

### Table 3.2. MSE vs. Time series length

<table>
<thead>
<tr>
<th>n</th>
<th>MSE($\hat{A}_{n,m}$)</th>
<th>MSE($\hat{B}_{n,m}$)</th>
<th>MSE($\hat{\omega}_{n,m}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>1.497023000</td>
<td>2.070980000</td>
<td>0.000001043</td>
</tr>
<tr>
<td>200</td>
<td>0.391485800</td>
<td>0.476013900</td>
<td>0.000001349</td>
</tr>
<tr>
<td>300</td>
<td>0.252940200</td>
<td>0.283450400</td>
<td>0.000000366</td>
</tr>
<tr>
<td>400</td>
<td>0.172374900</td>
<td>0.205647500</td>
<td>0.000000148</td>
</tr>
<tr>
<td>500</td>
<td>0.136898900</td>
<td>0.151340500</td>
<td>0.000000074</td>
</tr>
<tr>
<td>600</td>
<td>0.107044800</td>
<td>0.124402200</td>
<td>0.000000039</td>
</tr>
<tr>
<td>700</td>
<td>0.098976110</td>
<td>0.104186700</td>
<td>0.000000028</td>
</tr>
<tr>
<td>800</td>
<td>0.089476890</td>
<td>0.084489880</td>
<td>0.000000019</td>
</tr>
<tr>
<td>900</td>
<td>0.074666720</td>
<td>0.086571040</td>
<td>0.000000013</td>
</tr>
<tr>
<td>1000</td>
<td>0.060472530</td>
<td>0.071257020</td>
<td>0.000000009</td>
</tr>
</tbody>
</table>

### Table 3.3. Bias vs. Time series length

<table>
<thead>
<tr>
<th>n</th>
<th>Bias($\hat{A}_{n,m}$)</th>
<th>Bias($\hat{B}_{n,m}$)</th>
<th>Bias($\hat{\omega}_{n,m}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>2.085070e-01</td>
<td>3.402365e-01</td>
<td>-8.554667e-05</td>
</tr>
<tr>
<td>200</td>
<td>3.532309e-02</td>
<td>1.313686e-01</td>
<td>3.057407e-05</td>
</tr>
<tr>
<td>300</td>
<td>4.287946e-02</td>
<td>7.585903e-02</td>
<td>-1.376460e-05</td>
</tr>
<tr>
<td>400</td>
<td>4.383867e-02</td>
<td>8.411713e-02</td>
<td>5.294376e-06</td>
</tr>
<tr>
<td>500</td>
<td>5.475575e-02</td>
<td>6.497654e-02</td>
<td>-1.055630e-05</td>
</tr>
<tr>
<td>600</td>
<td>2.421265e-02</td>
<td>3.037480e-02</td>
<td>3.932200e-07</td>
</tr>
<tr>
<td>700</td>
<td>1.900718e-02</td>
<td>4.304497e-02</td>
<td>4.795330e-06</td>
</tr>
<tr>
<td>800</td>
<td>1.807705e-02</td>
<td>3.410545e-02</td>
<td>1.125110e-06</td>
</tr>
<tr>
<td>900</td>
<td>1.797766e-02</td>
<td>3.397651e-02</td>
<td>7.082500e-07</td>
</tr>
<tr>
<td>1000</td>
<td>1.512700e-02</td>
<td>3.257850e-02</td>
<td>-2.343000e-06</td>
</tr>
</tbody>
</table>

### Table 3.4. Means vs. Time series length

<table>
<thead>
<tr>
<th>n</th>
<th>mean($\hat{A}_{n,m}$)</th>
<th>mean($\hat{B}_{n,m}$)</th>
<th>mean($\hat{\omega}_{n,m}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>2.208507000</td>
<td>4.340237000</td>
<td>0.399914500</td>
</tr>
<tr>
<td>200</td>
<td>2.035323000</td>
<td>4.131369000</td>
<td>0.400030600</td>
</tr>
<tr>
<td>300</td>
<td>2.042879000</td>
<td>4.075859000</td>
<td>0.399986200</td>
</tr>
<tr>
<td>400</td>
<td>2.043839000</td>
<td>4.084117000</td>
<td>0.400005300</td>
</tr>
<tr>
<td>500</td>
<td>2.054756000</td>
<td>4.064977000</td>
<td>0.399989400</td>
</tr>
<tr>
<td>600</td>
<td>2.024130000</td>
<td>4.050375000</td>
<td>0.400004400</td>
</tr>
<tr>
<td>700</td>
<td>2.010908000</td>
<td>4.043045000</td>
<td>0.400004800</td>
</tr>
<tr>
<td>800</td>
<td>2.018077000</td>
<td>4.034105000</td>
<td>0.400001100</td>
</tr>
<tr>
<td>900</td>
<td>2.017978000</td>
<td>4.033977000</td>
<td>0.400007000</td>
</tr>
<tr>
<td>1000</td>
<td>2.015127000</td>
<td>4.032578000</td>
<td>0.399997700</td>
</tr>
</tbody>
</table>
**Table 3.5. KS test for normality: p-values**

<table>
<thead>
<tr>
<th>n</th>
<th>KS.p.(A_{n,m})</th>
<th>KS.p.(B_{n,m})</th>
<th>KS.p(\hat{\omega}_{n,m})</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0.000</td>
<td>0.000</td>
<td>0.774</td>
</tr>
<tr>
<td>200</td>
<td>0.372</td>
<td>0.073</td>
<td>0.646</td>
</tr>
<tr>
<td>300</td>
<td>0.588</td>
<td>0.094</td>
<td>0.801</td>
</tr>
<tr>
<td>400</td>
<td>0.838</td>
<td>0.012</td>
<td>0.711</td>
</tr>
<tr>
<td>500</td>
<td>0.559</td>
<td>0.433</td>
<td>0.676</td>
</tr>
<tr>
<td>600</td>
<td>0.305</td>
<td>0.072</td>
<td>0.973</td>
</tr>
<tr>
<td>700</td>
<td>0.697</td>
<td>0.081</td>
<td>0.761</td>
</tr>
<tr>
<td>800</td>
<td>0.949</td>
<td>0.413</td>
<td>0.963</td>
</tr>
<tr>
<td>900</td>
<td>0.957</td>
<td>0.431</td>
<td>0.851</td>
</tr>
<tr>
<td>1000</td>
<td>0.417</td>
<td>0.614</td>
<td>0.715</td>
</tr>
</tbody>
</table>

**Table 3.6. Shapiro-Wilks test: p-values**

<table>
<thead>
<tr>
<th>n</th>
<th>SW.p.(A_{n,m})</th>
<th>SW.p.(B_{n,m})</th>
<th>SW.p(\hat{\omega}_{n,m})</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0.000</td>
<td>0.000</td>
<td>0.132</td>
</tr>
<tr>
<td>200</td>
<td>0.004</td>
<td>0.000</td>
<td>0.150</td>
</tr>
<tr>
<td>300</td>
<td>0.002</td>
<td>0.000</td>
<td>0.411</td>
</tr>
<tr>
<td>400</td>
<td>0.771</td>
<td>0.000</td>
<td>0.450</td>
</tr>
<tr>
<td>500</td>
<td>0.149</td>
<td>0.000</td>
<td>0.096</td>
</tr>
<tr>
<td>600</td>
<td>0.043</td>
<td>0.000</td>
<td>0.919</td>
</tr>
<tr>
<td>700</td>
<td>0.289</td>
<td>0.000</td>
<td>0.857</td>
</tr>
<tr>
<td>800</td>
<td>0.335</td>
<td>0.000</td>
<td>0.497</td>
</tr>
<tr>
<td>900</td>
<td>0.981</td>
<td>0.001</td>
<td>0.775</td>
</tr>
<tr>
<td>1000</td>
<td>0.030</td>
<td>0.011</td>
<td>0.170</td>
</tr>
</tbody>
</table>
APPENDIX

USEFUL THEOREMS AND PROOFS OF INTERMEDIATE RESULTS
APPENDIX

USEFUL THEOREMS AND PROOFS OF INTERMEDIATE RESULTS

1. Proofs of Lemmas & Corollaries

1.1. Proof of Lemma 1

PROOF. The proof uses power series expansion of $f$ and the Cauchy- Hadamard formula. $f(X(t)) = \sum_{m=0}^{\infty} a_m X(t)^m$, where $\limsup_{m\to\infty} |a_m|^{1/m} = 0$. So $\exists M \in \mathbb{N}$ such that $\forall m \geq M, |a_m X(t)^m| < (1/2)^m$, since $X$ is bounded. And $\sum_{m=M+1}^{\infty} |a_m X(t)^m| < (1/2)^{M+1} + (1/2)^{M+2} + \ldots = (1/2)^M$. Let $M(\delta) := \max \{M, \log_2(1/\delta)\}$. Then

$$\frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k f(X(t)) = \frac{1}{n} \sum_{t=1}^{n} \left(\frac{t}{n}\right)^k f(X(t))$$

$$= \frac{1}{n} \sum_{t=1}^{n} \left(\frac{t}{n}\right)^k \left(\sum_{m=0}^{M(\delta)} a_m X(t)^m + \sum_{m=M(\delta)+1}^{\infty} a_m X(t)^m\right)$$

$$= \frac{1}{n} \sum_{t=1}^{n} \left(\frac{t}{n}\right)^k \sum_{m=0}^{M(\delta)} a_m X(t)^m + \frac{1}{n} \sum_{t=1}^{n} \left(\frac{t}{n}\right)^k \sum_{m=M(\delta)+1}^{\infty} a_m X(t)^m$$

Noting that

$$\frac{1}{n} \sum_{t=1}^{n} \left(\frac{t}{n}\right)^k \sum_{m=M(\delta)+1}^{\infty} a_m X(t)^m \leq \frac{1}{n} \sum_{t=1}^{n} \left(\frac{t}{n}\right)^k \sum_{m=M(\delta)+1}^{\infty} |a_m X(t)^m| < \delta$$

completes proof of the claim. \qed

1.2. Proof of Lemma 2

PROOF. Use induction.

Basis step: $j=k=0$: Note that we may use the following identities (mentioned in Walker [11])

$$\sum_{t=1}^{n} \exp(i\omega t) = e^{i(n+1)\omega/2} \sin(n\omega/2)/(\sin(\omega/2)) = O(1), \text{ if } \omega \in (0, 2\pi)$$

$$\sum_{t=1}^{n} \exp(i\omega t) = n, \text{ if } \omega = 0 \text{ or } 2\pi$$

31
Inductive step: Suppose \( P(j, k, \omega) = b_{jk}n + O(1) \) and \( Q(j, k, \omega) = c_{jk}n + O(1) \), for some \( b_{jk}, c_{jk} \in \mathbb{R} \).

Then using the product-to-sum trigonometric identities, one can show that

\[
\begin{align*}
P(j + 1, k, \omega) &= \frac{1}{2}P(j, k, \omega + \omega_0) + \frac{1}{2}P(j, k, \omega - \omega_0) \\
P(j, k + 1, \omega) &= \frac{1}{2}Q(j, k, \omega + \omega_0) - \frac{1}{2}Q(j, k, \omega - \omega_0) \\
Q(j + 1, k, \omega) &= \frac{1}{2}Q(j, k, \omega + \omega_0) + \frac{1}{2}Q(j, k, \omega - \omega_0) \\
Q(j, k + 1, \omega) &= \frac{1}{2}P(j, k, \omega - \omega_0) - \frac{1}{2}P(j, k, \omega + \omega_0)
\end{align*}
\]

In each case, the order identity is verified. \( \square \)

1.3. Proof of Lemma 3

**Proof.** Since

\[
\frac{1}{n} \sum_{t=1}^{n} \sum_{m=0}^{M} a_m X(t)^m = \sum_{m=0}^{M} a_m \frac{1}{n} \sum_{t=1}^{n} X(t)^m
\]

It follows that for as long as \( X(t) \) is bounded uniformly in \( t \), we may use the argument in the proof of Lemma 1, which essentially satisfies the Cauchy criterion, to prove that the following exists for any \( n \).

\[
\sum_{m=0}^{\infty} a_m \frac{1}{n} \sum_{t=1}^{n} (X(t))^m
\]

Hence, by the uniqueness of limits of Cauchy sequences,

\[
\frac{1}{n} \sum_{t=1}^{n} \sum_{m=0}^{\infty} a_m (X(t))^m = \sum_{m=0}^{\infty} a_m \frac{1}{n} \sum_{t=1}^{n} (X(t))^m
\]

\( \square \)
1.4. Proof of Corollary 2

**Proof.** Note that

$$\lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} \sum_{m=0}^{M} a_m X(t)^m = \sum_{m=0}^{M} \lim_{n \to \infty} \sum_{t=1}^{n} \frac{1}{n} X(t)^m$$

As \( m \to \infty \), the right hand side of the above converges to

$$\sum_{m=0}^{\infty} a_m \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} X(t)^m$$

which exists by a verification of the Cauchy criterion as similarly done in the proof of Lemma 1. □

1.5. Proof of Lemma 4

**Proof.** We’ll prove that the main diagonal entries in the matrix \( U \) are positive, as well as that the factors of the determinant are all positive. The key lies in the observation that

$$p_{0i}(1 - p_{0i}) = \frac{\exp(f(\theta_0, t))}{(1 + \exp(f(\theta_0, t)))^2} \geq c := \frac{\exp(|A_0| + |B_0|)}{(1 + \exp(|A_0| + |B_0|))^2}$$

So then \( d_1 \geq \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} c (\cos \omega_0 t)^2 = c/2 > 0 \). Similarly, \( d_2 > 0 \).

$$d_1 d_2 - d_3^2 = \lim_{n \to \infty} \frac{1}{2n^2} \sum_{i=1}^{n} p_{0i}(1 - p_{0i}) \sum_{j=1}^{n} p_{0j}(1 - p_{0j}) (\cos \omega_0 i \sin \omega_0 j - \sin \omega_0 i \cos \omega_0 j)^2$$

Again, based on the observation that \( p_{0i}(1 - p_{0i}) \) is bounded below uniformly in \( t \), the quantity is bounded below by

$$\lim_{n \to \infty} \frac{1}{2n^2} \sum_{i=1}^{n} c^2 \sum_{j=1}^{n} (\cos \omega_0 i \sin \omega_0 j - \sin \omega_0 i \cos \omega_0 j)^2$$

which is equal to \( c^2(X + Y + Z) \), where

$$X := \lim_{n \to \infty} \frac{1}{2n^2} \sum_{i=1}^{n} (\cos \omega_0 i)^2 \sum_{j=1}^{n} (\sin \omega_0 j)^2$$

$$Y := \lim_{n \to \infty} - \frac{2}{2n^2} \sum_{i=1}^{n} \sin(2\omega_0 i) \sum_{j=1}^{n} \sin(2\omega_0 j)$$

$$Z := \lim_{n \to \infty} \frac{1}{2n^2} \sum_{i=1}^{n} (\sin \omega_0 i)^2 \sum_{j=1}^{n} (\cos \omega_0 j)^2$$
Noting that \( X + Y + Z = \lim_{n \to \infty} \left( \frac{1}{8} (1 + O(n^{-1})) + O(n^{-1}) + \frac{1}{8} (1 + O(n^{-1})) \right) \) gives the limit as \( c^2/4 > 0 \).

On the other hand,

\[
A_0^2 d_2 + B_0^2 d_1 - 2A_0 B_0 d_3 = \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} p_{0t} (1 - p_{0t}) (A_0 \sin \omega_0 t - B_0 \cos \omega_0 t)^2
\]

\[
\geq \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} c (A_0 \sin \omega_0 t - B_0 \cos \omega_0 t)^2
\]

which is equal to

\[
\lim_{n \to \infty} c \left( A_0^2 (1/2 + O(n^{-1})) + B_0^2 (1/2 + O(n^{-1})) + O(n^{-1}) \right) = c(A_0^2 + B_0^2)/2
\]

This concludes the proof of the lemma. \(\square\)

1.6. Proof of Modified Lemma 5

**Proof.** The more general form of this lemma is the following: Let \( \hat{\theta}_n \) minimize \( F_n(\theta) \). Then if for any fixed \( \delta > 0 \), the following display holds:

\[
\inf_{||\theta - \theta_0|| \geq \delta} (F_n(\theta) - F_n(\theta_0)) = W_n(\delta) + K_n(\delta)
\]

then \( ||\hat{\theta}_n - \theta_0|| = o_p(1) \).

To prove this, let \( \delta > 0 \) be given. Suppose \( E_n(\delta) := \inf_{||\theta - \theta_0|| \geq \delta} (F_n(\theta) - F_n(\theta_0)) = W_n(\delta) + K_n(\delta) \), where \( K_n(\delta) \) is a deterministic quantity bounded below by a number \( K \in \mathbb{R}^+ \) uniformly in \( n \) (while holding the value of \( \delta \) fixed), and \( W_n(\delta) \) converges to 0 in probability uniformly in \( \theta \) (while holding the value of \( \delta \) fixed).

Let \( \epsilon > 0 \) be given. Without loss of generality assume \( \epsilon < K \). Now since \( W_n(\delta) \) converges to 0 in probability uniformly in \( \theta \),

\[
P(|E_n(\delta) - K_n(\delta)| > \epsilon) \to 0
\]

\[
P(|E_n(\delta) - K_n(\delta)| \leq \epsilon) \to 1
\]

\[
P(E_n(\delta) \geq K_n(\delta) - \epsilon) \to 1
\]

\[
P(E_n(\delta) \geq K - \epsilon) \to 1
\]
On the other hand, since \( F_n(\hat{\theta}_n) - F_n(\theta_0) \leq 0 \), we have that
\[
\mathbb{P}(|\hat{\theta}_n - \theta_0| > \delta) \leq \mathbb{P}(E_n(\delta) < K - \epsilon) \to 0
\]

Since \( \delta \) was arbitrary, this shows that \( \hat{\theta}_n \to_{P} \theta_0 \).

\[\square\]

1.7. Proof of Lemma 6

**Proof.** Based on the definition of \( \frac{1}{n} \{G_n(\theta) - G_n(\theta_0)\} \), it suffices to show that

\[
\sup_{\theta} \left| \frac{1}{n} \sum_{t=1}^{n} (y_t - p_{0t}) (f(\theta, t) - f(\theta_0, t)) \right| \to 0
\]

in probability (uniformly in \( \theta \)). Song & Li [8] noted (17)'s connection to the periodogram. We can mimic the proof in Walker [11] and Song & Li [8]:

Let \( \epsilon_t := (y_t - p_{0t}) \). Then \( \left| \sum_{t=1}^{n} (y_t - p_{0t}) (f(\theta, t) - f(\theta_0, t)) \right| \) is upper bounded by \( \left| \sum_{t=1}^{n} \epsilon_t e^{i\omega t} \right| \) times a weighted sum of \(|A|, |B|, |A_0|, |B_0|\), since \( f(\theta, t) = A \cos(\omega t) + B \sin(\omega t) \). We will use the fact that \( E(\epsilon_t) = 0, \epsilon_i, \epsilon_j \) independent \( \forall i \neq j \), and \( E(\epsilon_t^2) = p_{0t}(1 - p_{0t}) \leq 1/4 \) in showing that

\[
\sup_{\theta} \left| \sum_{t=1}^{n} (y_t - p_{0t}) (f(\theta, t) - f(\theta_0, t)) \right| = O_p(n^{3/4})
\]

from which (17) follows. To prove (18), note that

\[
\left| \sum_{t=1}^{n} \epsilon_t e^{i\omega t} \right|^2 = \sum_{t=1}^{n} \left| \epsilon_t e^{i\omega t} \cdot \sum_{s=1}^{n} \epsilon_s e^{-i\omega s} \right|^2
\]

\[
= \sum_{|s|=0}^{n-1} \sum_{t=1}^{n} \epsilon_t \epsilon_{t+|s|} \leq \sum_{|s|=0}^{n-1} \sum_{t=1}^{n-|s|} \sum_{|s|=0}^{n-|s|} \epsilon_t \epsilon_{t+|s|}
\]

The expectation of the last expression above is no more than (by separating \(|s| = 0\) from \(|s| > 0\) that of the following

\[
\sum_{t=1}^{n} \epsilon_t^2 + 2 \sum_{s=1}^{n-1} \sum_{t=1}^{n-s} \epsilon_t \epsilon_{t+s}
\]
whose first term’s expectation is no more than \( \frac{1}{4}n \), while with an application of Jensen’s inequality using the fact that the square function is convex, in particular, that

\[
\left( E \left| \sum_{t=1}^{n-s} \epsilon_t \epsilon_{t+s} \right| \right)^2 \leq E \left( \sum_{t=1}^{n-s} \epsilon_t \epsilon_{t+s} \right)^2 = E \left( \sum_{t=1}^{n-s} \epsilon_t \epsilon_{t+s} \right)^2
\]

Therefore, the expectation of \( 2 \sum_{s=1}^{n-1} \left| \sum_{t=1}^{n-s} \epsilon_t \epsilon_{t+s} \right| \) is no more than

\[
2 \sum_{s=1}^{n-1} \left( E \left( \sum_{t=1}^{n-s} \epsilon_t \epsilon_{t+s} \right)^2 \right)^{1/2}
\]

\[
= 2 \sum_{s=1}^{n-1} \left( \sum_{1 \leq i, j \leq n-s} E(\epsilon_i \epsilon_{i+s} \epsilon_j \epsilon_{j+s}) \right)^{1/2}
\]

\[
= 2 \sum_{s=1}^{n-1} \left( \sum_{t=1}^{n-s} E(\epsilon_t^2) \cdot E(\epsilon_{t+s}^2) \right)^{1/2}
\]

\[
\leq 2 \sum_{s=1}^{n-1} \frac{1}{4} \sqrt{n - s}
\]

The second equality due to the fact that \( E(\epsilon_i \epsilon_{i+s} \epsilon_j \epsilon_{j+s}) = 0 \) whenever \( i \neq j \): at most two of \( i, i+s, j, \) and \( j+s \) will be the same, and the rest will be different between themselves. So the expected value of their product in the case that \( i \neq j \) will be 0 based on the assumption that \( \epsilon_t \) are independent. On the other hand, when \( i = j, \epsilon_i \epsilon_{i+s} \epsilon_j \epsilon_{j+s} = \epsilon_i^2 \epsilon_{i+s}^2 \). Since the last expression in the above display is no more than \( \frac{1}{2} n^{3/2} \), (18) holds.

1.8. Proof of Lemma 7

**Proof.** In particular, we would like to show that

\[
U_n(\theta_0 + s(\hat{\theta}_n - \theta_0)) - U_n(\theta_0) = o_p(1) \quad \text{uniformly in } s \in [0, 1]
\]

Assuming the consistency \( \hat{\theta}_n^* = o_p(1) \), and observing that the deterministic entries of the matrix \( U_n(\hat{\theta}_n) \) are all of the form

\[
\frac{1}{n^{k+1}} \sum_{t=1}^{n} p_t (1 - p_t) t^k g(\theta, t)
\]

where \( p_t \) and \( g(\theta, t) \) are both evaluated at \( \theta = \hat{\theta}_n \). the proof is divided in the following two steps:
where the expression for \( g(22) \) is

\[
U_n^*(\theta_0 + s(\hat{\theta}_n - \theta_0)) - U_n^*(\theta_0) = o_p(1) \quad \text{uniformly in } s \in [0, 1].
\]

(21)

\[
V_n(\theta_0 + s(\hat{\theta}_n - \theta_0)) - V_n(\theta_0) = o_p(1) \quad \text{uniformly in } s \in [0, 1].
\]

Proof of (20). If one sets \( \tilde{\theta}_n = \theta_0 + s(\hat{\theta}_n - \theta_0) \), then each entry in the difference \( U_n(\theta_0 + s(\hat{\theta}_n - \theta_0)) - U_n(\theta_0) \) has the form of

\[
\frac{1}{n^{k+1}} \sum_{t=1}^{n} p_t(1 - p_t) t^k g(\theta, t) - \frac{1}{n^{k+1}} \sum_{t=1}^{n} p_{0t}(1 - p_{0t}) t^k g(\theta_0, t),
\]

where the expression for \( g(\tilde{\theta}_n, t) \) depends on where the entry is in the matrix. For example,

\[
g(\theta, t) = (-A \sin \omega t + B \cos \omega t) \cos \omega t \text{ for the entry in the first row and third column of } U_n^*(\theta).
\]

Hence, to prove (20), it remains to show the following

(22)

\[
\frac{1}{n^{k+1}} \sum_{t=1}^{n} p_t(1 - p_t) t^k g(\theta, t) - \frac{1}{n^{k+1}} \sum_{t=1}^{n} p_{0t}(1 - p_{0t}) t^k g(\theta_0, t) = o_p(1).
\]

Note that

\[
\frac{1}{n^{k+1}} \sum_{t=1}^{n} p_t(1 - p_t) t^k g(\tilde{\theta}_n, t) - \frac{1}{n^{k+1}} \sum_{t=1}^{n} p_{0t}(1 - p_{0t}) t^k g(\theta_0, t)
\]

\[
= \frac{1}{n^{k+1}} \sum_{t=1}^{n} p_t(1 - p_t) t^k (g(\tilde{\theta}_n, t) - g(\theta_0, t)) + \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k (\tilde{p}_t(1 - p_t) - p_{0t}(1 - p_{0t})) g(\theta_0, t)
\]

\[
= \frac{1}{n^{k+1}} \sum_{t=1}^{n} p_t(1 - p_t) t^k (g(\tilde{\theta}_n, t) - g(\theta_0, t)) + \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k g(\theta_0, t) \sum_{m=0}^{\infty} a_m ((\tilde{\theta}_n, t)^m - \theta_0(t)^m)
\]

where \( \{a_m\} \) are the coefficients in the power series expansion of \( \frac{e^x}{(1+e^x)^2} \) around \( x = 0 \). We will prove the uniform convergence in probability for the lower right entry in \( U_n(\tilde{\theta}_n) - U_n(\theta_0) \).

The proofs for the other entries are similar. For the lower right entry,

\[
g(\tilde{\theta}_n, t) = (-\tilde{A}_n \sin \tilde{\omega}_n t + \tilde{B}_n \cos \tilde{\omega}_n t)^2
\]

Then it follows that

\[
g(\tilde{\theta}_n, t) - g(\theta_0, t) = \tilde{B}_n^2((\cos \tilde{\omega}_n t)^2 - (\cos \omega_0 t)^2) + \tilde{A}_n^2((\sin \tilde{\omega}_n t)^2 - (\sin \omega_0 t)^2)
\]

\[
- \tilde{A}_n \tilde{B}_n (\sin 2\tilde{\omega}_n t - \sin 2\omega_0 t) + \sin \omega_0 t (\tilde{A}_n^2 - A_0^2) + (\cos \omega_0 t)^2 (\tilde{B}_n^2 - B_0^2)
\]

\[
- \sin(2\omega_0 t) (\tilde{A}_n \tilde{B}_n - A_0 B_0)
\]

37
\[ \pm \tilde{B}_n^2(\cos \tilde{\omega}_n t + \cos \omega_0 t)(\cos \tilde{\omega}_n t - \cos \omega_0 t) + \tilde{A}_n(\sin \tilde{\omega}_n t + \sin \omega_0 t)(\sin \omega_n t - \sin \omega_0 t) \]
\[ - \tilde{A}_n \tilde{B}_n(\sin 2\tilde{\omega}_n t - \sin 2\omega_0 t) + \sin^2 \omega_0 t(\tilde{A}_n + A_0)(\tilde{A}_n - A_0) \]
\[ + \cos^2 \omega_0 t(\tilde{B}_n + B_0)(\tilde{B}_n - B_0) - \sin(2\omega_0 t)(\tilde{A}_n(\tilde{B}_n - B_0) + B_0(\tilde{A}_n - A_0)) \]

where each term is \( o_p(1) \) uniformly in \( t \) and \( \tilde{\theta}_n \). Using Taylor series expansions and mean value theorem, there exist for each \( t \), a \( \omega_t^* \) and \( \omega_t^{**} \) such that the following hold:

\[ |\cos \tilde{\omega}_n t - \cos \omega_0 t| = |t(\sin \omega_t^* t)(\tilde{\omega}_n - \omega_0)| \leq n|\tilde{\omega}_n - \omega_0| \]
\[ |\sin k\tilde{\omega}_n t - \sin k\omega_0 t| = |k t(\cos \omega_t^{**} t)(\tilde{\omega}_n - \omega_0)| \leq k n|\tilde{\omega}_n - \omega_0| \]
\[ |\tilde{A}_n \tilde{B}_n - A_0 B_0| = |\tilde{A}_n(\tilde{B}_n - B_0) + B_0(\tilde{A}_n - A_0)| \]

which are all \( o_p(1) \) for finite \( k \in \mathbb{N} \) and uniformly in \( t \) and \( \tilde{\theta}_n \) due to the fact that \( \tilde{\theta}_n^* = o_p(1) \). Moreover, since \( g(\tilde{\theta}_n, t) - g(\theta_0, t) \) is a sum of finitely many terms, each of which in absolute value is bounded above by the product of a uniform constant \( K := \sup_{(A, B)} \{1, |A| + |B|\} \) (due to the fact that \( A, B \) attain values in a compact set) and one of \( |\tilde{A}_n - A_0|, |\tilde{B}_n - B_0|, \) and \( |n(\tilde{\omega}_n - \omega_0)| \). The last three quantities are all \( o_p(1) \) due to consistency of \( \tilde{\theta}_n^* \). Therefore, we may conclude that uniformly in \( t \) and \( \tilde{\theta}_n \),

\[ |g(\tilde{\theta}_n, t) - g(\theta_0, t)| \leq 2K^2(|\tilde{A}_n - A_0| + |\tilde{B}_n - B_0| + |n(\tilde{\omega}_n - \omega_0)|) = o_p(1) \]

since

\[ 2K^2(|\tilde{A}_n - A_0| + |\tilde{B}_n - B_0| + |n(\tilde{\omega}_n - \omega_0)|) \leq 2K^2(|\tilde{A}_n - A_0| + |\tilde{B}_n - B_0| + |n(\tilde{\omega}_n - \omega_0)|) \]

Let

\[ K := \sup_{(A, B)} \{1, |A| + |B|\} \]

Then \( K \) is bounded because \( A, B \) attain values in a compact set by assumption.

\[ \left| \frac{1}{n^{k+1}} \sum_{t=1}^{n} \tilde{p}_t(1 - \tilde{p}_t) t^k (g(\tilde{\theta}_n, t) - g(\theta_0, t)) \right| \]
\[
\begin{align*}
&= \left| \frac{1}{n^{k+1}} \sum_{t=1}^{n} \sum_{m=0}^{\infty} a_m f(\theta_0, t)^m t^k (g(\tilde{\theta}_n, t) - g(\theta_0, t)) \right| \\
&= \left| \sum_{m=0}^{\infty} a_m \cdot \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k f(\theta_0, t)^m (g(\tilde{\theta}_n, t) - g(\theta_0, t)) \right| \\
&\leq \sum_{m=0}^{\infty} |a_m| \cdot \frac{1}{n^{k+1}} \left\{ \left( \sum_{t=1}^{n} t^{2k} \cdot K^{2m} \right) \left( \sum_{t=1}^{n} (g(\tilde{\theta}_n, t) - g(\theta_0, t))^2 \right) \right\}^{1/2} \\
&\leq \sum_{m=0}^{\infty} |a_m| \cdot \frac{1}{n^{k+1}} \left\{ K^{2m} \cdot (n^{2k+1}) \cdot \sum_{t=1}^{n} 4K^4 (|\tilde{\theta}_n - A_0| + |\tilde{B}_n - B_0| + |n(\tilde{\omega}_n - \omega_0)|)^2 \right\}^{1/2} \\
&= \sum_{m=0}^{\infty} |a_m| \cdot \frac{1}{n^{k+1}} \left( K^{2m} \cdot (n^{2k+1}) \cdot n \cdot 4K^4 (|\tilde{\theta}_n - A_0| + |\tilde{B}_n - B_0| + |n(\tilde{\omega}_n - \omega_0)|)^2 \right)^{1/2} \\
&= 2K^2 (|\tilde{\theta}_n - A_0| + |\tilde{B}_n - B_0| + |n(\tilde{\omega}_n - \omega_0)|) \sum_{m=0}^{\infty} |a_m| \cdot K^m \\
&= o_p(1) \sum_{m=0}^{\infty} |a_m| \cdot K^m
\end{align*}
\]

The second equality above follows from (7), and the first inequality follows from Cauchy-Schwarz. Now that \( \limsup_{m \to \infty} (|a_m|)^{1/m} = 0 \), so \( \sum_{m=0}^{\infty} |a_m| \cdot K^M < \infty \) exists. Moreover, the \( o_p(1) \) is uniform from the display (24). Therefore, uniformly in \( \tilde{\theta}_n = \theta_0 + s(\tilde{\theta}_n - \theta_0) \), \( s \in [0, 1] \),

\begin{equation}
\begin{align*}
&= \sum_{m=0}^{\infty} \frac{1}{n^{k+1}} \sum_{t=1}^{n} \tilde{p}_t (1 - \tilde{p}_t) t^k (g(\tilde{\theta}_n, t) - g(\theta_0, t)) = o_p(1)
\end{align*}
\end{equation}

On the other hand,

\[
(f(\tilde{\theta}_n, t))^m - (f(\theta_0, t))^m = (f(\tilde{\theta}_n, t) - f(\theta_0, t)) S
\]

where \( S = (f(\tilde{\theta}_n, t))^{m-1} + (f(\tilde{\theta}_n, t))^{m-2} f(\theta_0, t) + \ldots + (f(\theta_0, t))^{m-1} \).

and \( f(\tilde{\theta}_n, t) - f(\theta_0, t) = \tilde{A}_n (\cos \tilde{\omega}_n t - \cos \omega_0 t) + \tilde{B}_n (\sin \tilde{\omega}_n t - \sin \omega_0 t) + \cos \omega_0 t (\tilde{A}_n - A_0) + \sin \omega_0 t (\tilde{B}_n - B_0) \)

In fact, \( S \leq mK^{m-1} \) uniformly in \( t \) and \( \tilde{\theta}_n \). In addition, \( f(\tilde{\theta}_n, t) - f(\theta_0, t) = o_p(1) \) uniformly in \( t \) and \( \tilde{\theta}_n \) since \( \hat{\theta}_n^* = o_p(1) \) and \( |f(\tilde{\theta}_n, t) - f(\theta_0, t)| \) is bounded above by \( K(|n(\tilde{\omega}_n - \omega_0)| + \ldots \).
\[ |\tilde{A}_n - A_0| + |\tilde{B}_n - B_0| \]. Thus we have shown that uniformly in \( t \) and \( \tilde{\theta}_n \),

\[
(26) \quad |(f(\tilde{\theta}_n, t))^m - (f(\theta_0, t))^m| \leq mK^m(|n(\tilde{\omega}_n - \omega_0)| + |\tilde{A}_n - A_0| + |\tilde{B}_n - B_0|)
\]

Therefore,

\[
\left| \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k (\tilde{p}_t(1 - \tilde{p}_t) - p_0(1 - p_0)) g(\theta_0, t) \right|
\]

\[
= \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k \left( \sum_{m=0}^{\infty} a_m (f(\tilde{\theta}_n, t)^m - f(\theta_0, t)^m) g(\theta_0, t) \right)
\]

\[
\leq \sum_{m=0}^{\infty} |a_m| \frac{1}{n^{k+1}} \left( \sum_{t=1}^{n} t^{2k} \sum_{m=0}^{\infty} (f(\tilde{\theta}_n, t)^m - f(\theta_0, t)^m)^2 g(\theta_0, t)^2 \right)^{1/2}
\]

\[
\leq \sum_{m=0}^{\infty} |a_m| \frac{1}{n^{k+1}} \left( n^{2k+1} \cdot \sum_{t=1}^{n} \left( m^2 K^{2m}(|n(\tilde{\omega}_n - \omega_0)| + |\tilde{A}_n - A_0| + |\tilde{B}_n - B_0|)^2 \right) \right)^{1/2}
\]

\[
= \sum_{m=0}^{\infty} |a_m| \frac{1}{n^{k+1}} \left( n^{2k+2} m^2 K^{2m+2} (|n(\tilde{\omega}_n - \omega_0)| + |\tilde{A}_n - A_0| + |\tilde{B}_n - B_0|)^2 \right)^{1/2}
\]

\[
= (|n(\tilde{\omega}_n - \omega_0)| + |\tilde{A}_n - A_0| + |\tilde{B}_n - B_0|) \cdot \sum_{m=0}^{\infty} |a_m| \cdot mK^{m+1}
\]

Now since

\[
\limsup_{m \to \infty} (|a_m| mK)^{1/m} = \limsup_{m \to \infty} |a_m| \cdot 1 = 0
\]

It follows that \( \sum_{m=0}^{\infty} |a_m| \cdot mK^{m+1} < \infty \) exists and so the last quantity is \( o_p(1) \) uniformly in \( \tilde{\theta}_n \) due to (24). Thus, uniformly in \( t \) and \( \tilde{\theta}_n \),

\[
(27) \quad \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k (\tilde{p}_t(1 - \tilde{p}_t) - p_0(1 - p_0)) g(\theta_0, t) = o_p(1)
\]

Combining (25) and (27) we obtain (22). By performing similar analysis on the other entries of the matrix \( U_n \) one proves (20).

Proof of (21). To prove (21), namely, that

\[
V_n(\theta_0 + s(\tilde{\theta}_n - \theta_0)) - V_n(\theta_0) = o_p(1) \quad \text{uniformly in } s \in [0, 1].
\]
Observe that they are of the form

$$
\frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k (p_{0t} - \bar{p}_t) h_{i,j} (\bar{\theta}_n, t) - \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k (y_t - p_{0t}) h_{i,j} (\theta_0, t)
$$

where

$$
\begin{align*}
h_{i,j} (\bar{\theta}_n, t) &= \sin \bar{\omega}_n t, - \cos \bar{\omega}_n t, \quad \text{and} \quad \bar{A}_n \cos \bar{\omega}_n t + \bar{B}_n \sin \bar{\omega}_n t \\
\end{align*}
$$

respectively for the entries in the positions (1, 3), (2, 3), (3, 3), (3, 2), and (3, 1). In particular, Note that

$$
V_n (\theta_0 + s (\hat{\theta}_n - \theta_0)) = V_n^* (\theta_0 + s (\hat{\theta}_n - \theta_0)) + W_n (\theta_0 + s (\hat{\theta}_n - \theta_0))
$$

where

$$
\begin{align*}
V_n (\hat{\theta}_n) &:= \\
&= \begin{bmatrix}
0 & 0 & \frac{1}{n^2} \sum_{t=1}^{n} t^2 (y_t - \bar{p}_t) \sin \bar{\omega}_n t \\
\text{Symmetry} & 0 & -\frac{1}{n^2} \sum_{t=1}^{n} t^2 (y_t - \bar{p}_t) \cos \bar{\omega}_n t \\
& & \frac{1}{n^2} \sum_{t=1}^{n} t^2 (y_t - \bar{p}_t) (\bar{A}_n \cos \bar{\omega}_n t + \bar{B}_n \sin \bar{\omega}_n t)
\end{bmatrix}
\end{align*}
$$

$$
\begin{align*}
W_n (\hat{\theta}_n) &:= \\
&= \begin{bmatrix}
0 & 0 & \frac{1}{n^2} \sum_{t=1}^{n} t^2 (y_t - p_{0t}) \sin \bar{\omega}_n t \\
\text{Symmetry} & 0 & -\frac{1}{n^2} \sum_{t=1}^{n} t^2 (y_t - p_{0t}) \cos \bar{\omega}_n t \\
& & \frac{1}{n^2} \sum_{t=1}^{n} t^2 (y_t - p_{0t}) (\bar{A}_n \cos \bar{\omega}_n t + \bar{B}_n \sin \bar{\omega}_n t)
\end{bmatrix}
\end{align*}
$$

$$
\begin{align*}
V_n^* (\hat{\theta}_n) &:= \\
&= \begin{bmatrix}
0 & 0 & \frac{1}{n^2} \sum_{t=1}^{n} t^2 (p_{0t} - \bar{p}_t) \sin \bar{\omega}_n t \\
\text{Symmetry} & 0 & -\frac{1}{n^2} \sum_{t=1}^{n} t^2 (p_{0t} - \bar{p}_t) \cos \bar{\omega}_n t \\
& & \frac{1}{n^2} \sum_{t=1}^{n} t^2 (p_{0t} - \bar{p}_t) (\bar{A}_n \cos \bar{\omega}_n t + \bar{B}_n \sin \bar{\omega}_n t)
\end{bmatrix}
\end{align*}
$$

Note the following:

$$
\frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k (y_t - \bar{p}_t) h (\bar{\theta}_n, t) = \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k (y_t - p_{0t}) h (\theta_0, t) + \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k (p_{0t} - \bar{p}_t) h (\hat{\theta}_n, t)
$$

So in particular, if the following are both true uniformly in $\hat{\theta}_n$, then (21) is proved.

$$
\frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k (y_t - p_{0t}) h (\hat{\theta}_n, t) = o_p(1)
$$

$$
\frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k (p_{0t} - \bar{p}_t) h (\hat{\theta}_n, t) = o_p(1)
$$
The proof of (30) is based on Walker [11] and is similar to the proof of (17), by observing that the entries in $W_n(\tilde{\theta}_n)$ are weighted sums of $n^{-(k+1)} \sum_{t=1}^n t^k e^{it\tilde{\omega}_n} (y_t - p_{0t})$ with $O_p(1)$ weights [8], $k = 1, 2$. Hence according to Walker, they are $O_p(n^{-1/4})$ uniformly in $\tilde{\omega}_n$. For example, for the lower right entry in $W_n(\tilde{\theta}_n)$, its absolute value is

$$\left| \frac{1}{n^3} \sum_{t=1}^n t^2 (y_t - p_{0t}) (\tilde{A}_n \cos \tilde{\omega}_n t + \tilde{B}_n \sin \tilde{\omega}_n t) \right|$$

$$= \left| \tilde{A}_n \cdot \frac{1}{n^3} \sum_{t=1}^n t^2 (y_t - p_{0t}) \cos \tilde{\omega}_n t + \tilde{B}_n \cdot \frac{1}{n^3} \sum_{t=1}^n t^2 (y_t - p_{0t}) \sin \tilde{\omega}_n t \right|$$

$$\leq \left| \tilde{A}_n \right| \cdot \max_{0 \leq \tilde{\omega}_n \leq \pi} \left| \frac{1}{n^3} \sum_{t=1}^n t^2 (y_t - p_{0t}) \cos \tilde{\omega}_n t \right| + \left| \tilde{B}_n \right| \cdot \max_{0 \leq \tilde{\omega}_n \leq \pi} \left| \frac{1}{n^3} \sum_{t=1}^n t^2 (y_t - p_{0t}) \sin \tilde{\omega}_n t \right|$$

Let $\epsilon_t := y_t - p_{0t}$. Note that

$$\sqrt{\sum_{t=1}^n \epsilon_t^2 t^2 \cos \tilde{\omega}_n t}^2 + \sum_{t=1}^n \epsilon_t^2 t^2 \sin \tilde{\omega}_n t$$

$$= \left| \sum_{t=1}^n \epsilon_t^2 t^2 e^{i\tilde{\omega}_n t} \right|^2$$

$$= \sum_{t=1}^n \epsilon_t^2 t^2 \sum_{s=1}^n \epsilon_s^2 s^2 e^{-i\omega s}$$

$$= \sum_{|s| \leq n-1} \epsilon_s^2 s^2 \sum_{t=1}^{n-|s|} \epsilon_t^2 t^2 (t + |s|)^2$$

$$\leq \sum_{|s| \leq n-1} \epsilon_s^2 s^2 \sum_{t=1}^{n-|s|} \epsilon_t^2 t^2 (t + |s|)^2$$

The expectation of the last term above is no more than (by separating $|s| = 0$ from the case $|s| > 0$) that of the following:

$$\sum_{t=1}^n \epsilon_t^2 t^4 + 2 \sum_{s=1}^{n-1} \sum_{t=1}^{n-s} \epsilon_t^2 \epsilon_{t+s} t^2 (t + s)^2$$
On the other hand, using the fact that the square function is convex, then one can apply Jensen’s inequality to obtain that

\[
\left( E \left| \sum_{t=1}^{n-s} \epsilon_t \epsilon_{t+s} t^2 (t + s)^2 \right| \right)^2 \leq E \left( \sum_{t=1}^{n-s} \epsilon_t \epsilon_{t+s} t^2 (t + s)^2 \right)^2
\]

So then for \( s > 0 \),

\[
E \left| \sum_{t=1}^{n-s} \epsilon_t \epsilon_{t+s} t^2 (t + s)^2 \right| \leq \left\{ E \left( \sum_{t=1}^{n-s} \epsilon_t \epsilon_{t+s} t^2 (t + s)^2 \right)^2 \right\}^{1/2}
\]

\[
= \left\{ E \left( \sum_{1 \leq i,j \leq n-s} \epsilon_i \epsilon_{i+s} \epsilon_j \epsilon_{j+s} t^2 (i + s)^2 (j + s)^2 \right) \right\}^{1/2}
\]

\[
= \left\{ E \left( \sum_{t=1}^{n-s} \epsilon_t^2 E(\epsilon_{t+s}^2) t^4 (t + s)^4 \right) + E \left( \sum_{1 \leq i \neq j \leq n-s} \epsilon_i \epsilon_{i+s} \epsilon_j \epsilon_{j+s} t^2 (i + s)^2 (j + s)^2 \right) \right\}^{1/2}
\]

\[
= \left( \sum_{t=1}^{n-s} E(\epsilon_t^2) E(\epsilon_{t+s}^2) t^4 (t + s)^4 + 0 \right)^{1/2}
\]

\[
= \left( \sum_{t=1}^{n-s} \frac{1}{16} t^4 (t + s)^4 \right)^{1/2} \leq \left( \frac{1}{16} (n-s)^5 n^4 \right)^{1/2} = \frac{1}{4} (n-s)^{5/2} \cdot n^2
\]

Thus it follows that

\[
E \left\| \sum_{t=1}^{n} \epsilon_t t^2 e^{i\omega t} \right\|^2 \leq E \left( \sum_{t=1}^{n} \epsilon_t^2 t^4 \right) + 2(n-1) \cdot \frac{1}{4} (n-1)^{5/2} n^2
\]

\[
\leq \frac{1}{4} n^5 + \frac{1}{2} n^{11/2} \leq n^{11/2}
\]

Therefore,

\[
\left| \sum_{t=1}^{n} \epsilon_t t^2 \cos \tilde{\omega}_n t \right| \leq \max_{0 \leq \omega \leq \pi} \left( \left\| \sum_{t=1}^{n} \epsilon_t t^2 e^{i\omega t} \right\| \right)^{1/2} = O_p(n^{11/4})
\]

Similarly,

\[
\left| \sum_{t=1}^{n} \epsilon_t t^2 \sin \tilde{\omega}_n t \right| \leq \max_{0 \leq \omega \leq \pi} \left( \left\| \sum_{t=1}^{n} \epsilon_t t^2 e^{i\omega t} \right\| \right)^{1/2} = O_p(n^{11/4})
\]

Now that \(|\tilde{A}_n|, |\tilde{B}_n| \leq K < \infty\) uniformly because the estimates of \( a, B \) are in a compact set, it follows that the lower right entry of \( W_n(\tilde{\theta}_n) \) is \( O_p(n^{11/4}/n^3) = O_p(n^{-1/4}) \) uniformly in \( \tilde{\theta}_n \).
Utilizing the same argument above on the other entries of the matrix $W_n(\tilde{\theta}_n)$ shows that it is $O_p(n^{-1/4})$ uniformly in $\tilde{\theta}_n$. Namely,

$$W_n(\tilde{\theta}_n) \to P_0$$

uniformly in $\tilde{\theta}_n$.

The proof of (31) is similar to that of (25) and (27) in that the same power series expansion technique is used, with $\{b_m\}$ being the coefficient from the power series expansion of the function $p_t(u) := \frac{e^u}{1+e^u}$ with $u = f(\theta, t)$ and at around $u = 0$. Since $p_t(u)$ is real-analytic throughout the domain of $u$, Cauchy-Hadamard formula gives $\lim \sup_{m \to \infty} |b_m|^{1/m} = 0$. In particular,

$$\frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k (p_{0t} - \tilde{p}_t) h(\tilde{\theta}_n, t) = \frac{1}{n^{k+1}} \sum_{t=1}^{n} \sum_{m=0}^{\infty} b_m (f(\theta_0, t)^m - f(\tilde{\theta}_n, t)^m) h(\tilde{\theta}_n, t) = \sum_{m=0}^{\infty} b_m \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k (f(\theta_0, t)^m - f(\tilde{\theta}_n, t)^m) h(\tilde{\theta}_n, t)$$

The first equality following from the definition of $p_{0t}$ and $\tilde{p}_t$, and the second following from (7). To prove (31), note that by using (26) and that $h(\tilde{\theta}_n, t) = \tilde{A}_n \cos \tilde{\omega}_n t + \tilde{B}_n \sin \tilde{\omega}_n t$,

$$\left| \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k (p_{0t} - \tilde{p}_t) h(\tilde{\theta}_n, t) \right| \leq \sum_{m=0}^{\infty} |b_m| \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^{2m} \sum_{t=1}^{n} (f(\theta_0, t)^m - f(\tilde{\theta}_n, t)^m)^2 h(\tilde{\theta}_n, t)^2$$

$$\leq \sum_{m=0}^{\infty} |b_m| \frac{1}{n^{k+1}} \cdot \left( \sum_{t=1}^{n} t^{2m} K^2 (|n(\tilde{\omega}_n - \omega_0)| + |\tilde{A}_n - A_0| + |\tilde{B}_n - B_0|)^2 \cdot K^2 \right)^{1/2}$$

$$= \sum_{m=0}^{\infty} |b_m| m K^m (|n(\tilde{\omega}_n - \omega_0)| + |\tilde{A}_n - A_0| + |\tilde{B}_n - B_0|)$$

$$= (|n(\tilde{\omega}_n - \omega_0)| + |\tilde{A}_n - A_0| + |\tilde{B}_n - B_0|) \sum_{m=0}^{\infty} |b_m| m K^m$$
Now since
\[ \limsup_{m \to \infty} (|b_m| mK^{m+1})^{1/m} = \limsup_{m \to \infty} |b_m|^{1/m} \cdot K = 0 \]
It follows that \( \sum_{m=0}^{\infty} |b_m| \cdot mK^{m+1} < \infty \) exists and so the last quantity is \( o_p(1) \) due to (24). This finishes the proof of (31). Combining (31) and (30) gives (21), which when taken together with (20) proves (19).

2. Radchenko’s Lemma 1

\textbf{Radchenko’s Lemma 1.} Suppose that inequalities \( G_n(a_n, b_n) \leq G_n(0, 0) \) hold together with the stochastic bound \( \|(a_n, b_n)\| = o^*_p(1) \). Let \( \alpha \) and \( \beta \) be positive numbers satisfying \( \alpha \geq \beta \), and let \( \{\gamma_1, \ldots, \gamma_p, \eta_1, \ldots, \eta_p\} \) be a collection of nonnegative numbers satisfying \( \gamma_i < \alpha \) for all \( i \in \{1, \ldots, p\} \). Suppose that criterion functions \( G_n \) satisfy a representation

\[ G_n(a, b) - G_n(0, 0) = M_n(a, b) - N_n(a, b), \]

such that

\[ M_n(a_n, b_n) \geq ||a_n||^\alpha + ||b_n||^\beta \quad \text{with inner probability tending to one, and} \]

\[ [N_n(a_n, b_n)]^+ = O^*_p \left( \sum_{i \leq p} n^{-\eta_i} ||(a_n, b_n)||^{\gamma_i} \right). \]

Define \( \tau_n = \min_{i \leq p} \left( \frac{m_{\alpha - \gamma_i}}{\alpha - \gamma_i} \right) \). Then \( ||a_n|| = O^*_p \{n^{-\tau_0}\} \) and \( ||b_n|| = O^*_p \{n^{-\alpha \tau_0/\beta}\} \).

\textbf{Reference (Definitions and results concerning various quantities in the paper).}

\[ U_n(\tilde{\theta}_n) - U_n(\theta_0) \to_p 0 \]

\[ U_n(\tilde{\theta}) = D_n \nabla^2 G_n(\tilde{\theta}) D_n \]

\[ \tilde{\theta}_n := \theta_0 + s(\hat{\theta}_n - \theta_0) \]

\[ H_n(\theta_0 + R_n \alpha) = \int_0^1 (1 - s) \nabla^2 G_n(\theta_0 + s R_n \alpha) \, ds \]

\[ U_n(\tilde{\theta}_n)_{11} = n^{-1} \sum_{t=1}^{n} \hat{p}_t (1 - \hat{p}_t) (\cos \hat{\omega}_n t)^2 \]

\[ U_n(\tilde{\theta}_n)_{12} = n^{-1} \sum_{t=1}^{n} \hat{p}_t (1 - \hat{p}_t) \cos \hat{\omega}_n t \sin \hat{\omega}_n t \]
\[ U_n(\tilde{\theta}_n)_{13} = n^{-2} \sum_{t=1}^{n} \tilde{p}_t (1 - \tilde{p}_t) (-A \sin \tilde{\omega}_n t + B \cos \tilde{\omega}_n t) t \cos \tilde{\omega}_n t + t(y_t - \tilde{p}_t) \sin \tilde{\omega}_n t \]

\[ U_n(\tilde{\theta}_n)_{22} = n^{-1} \sum_{t=1}^{n} \tilde{p}_t (1 - \tilde{p}_t) (\sin \tilde{\omega}_n t)^2 \]

\[ U_n(\tilde{\theta}_n)_{23} = n^{-2} \sum_{t=1}^{n} \tilde{p}_t (1 - \tilde{p}_t) (-A \sin \tilde{\omega}_n t + B \cos \tilde{\omega}_n t) t \sin \tilde{\omega}_n t - t(y_t - \tilde{p}_t) \cos \tilde{\omega}_n t \]

\[ U_n(\tilde{\theta}_n)_{33} = n^{-3} \sum_{t=1}^{n} \tilde{p}_t (1 - \tilde{p}_t) (-A \sin \tilde{\omega}_n t + B \cos \tilde{\omega}_n t)^2 t^2 + t^2 (y_t - \tilde{p}_t) (A \cos \tilde{\omega}_n t + B \sin \tilde{\omega}_n t) \]

\[ Q_n(\theta_0 + R_n \alpha) = \alpha' D_n H_n(\theta_0 + R_n \alpha) D_n \alpha - \frac{1}{2} \alpha' U_n^*(\theta_0) \alpha \]

In particular,

\[ Q_n(\theta_0 + R_n \alpha) = \alpha' \int_{0}^{1} (1 - s) U_n(\theta_0 + s R_n \alpha) - U_n^*(\theta_0)) \, ds \alpha \]

\[ w_n(s, \alpha) := U_n(\theta_0 + s R_n \alpha) - U_n^*(\theta_0) \]


