Proceeding Paper

Asymptotic Distributions of M-Estimates for Parameters of Multivariate Time Series with Strong Mixing Property [†]

Alexander Kushnir * and Alexander Varypaev (ID)

Institute of Earthquake Prediction Theory and Mathematical Geophysics of Rassian Academy of Sciences, 113556 Moscow, Russia; avalex89@gmail.com
* Correspondence: afkushnir@gmail.com
† Presented at the 7th International conference on Time Series and Forecasting, Gran Canaria, Spain, 19–21 July 2021.

Abstract: The publication is devoted to studying asymptotic properties of statistical estimates of the distribution parameters $u \in R^q$ of a multidimensional random stationary time series $z_t \in R^m$, $t \in \mathbb{Z}$ satisfying the strong mixing conditions. We consider estimates $\hat{u}_n^\delta\left(\bar{z}_n\right)$, $\bar{z}_n = (z_1^T, \ldots, z_n^T)^T \in R^{mn}$ that provide in asymptotic $n \to \infty$ the maximum values for some objective functions $Q_n\left(\bar{z}_n; u\right)$, which have properties similar to the well-known property of local asymptotic normality. These estimates are constructed by solving the equations $\delta_n\left(\bar{z}_n; u\right) = 0$, where $\delta_n\left(\bar{z}_n; u\right)$ are arbitrary functions for which $\delta_n\left(\bar{z}_n; u\right) - \operatorname*{grad}_h Q_n\left(\bar{z}_n; u + n^{-1/2}h\right) \to 0$ $(n \to \infty)$ in $P_{n,u}\left(\bar{z}_n\right)$-probability uniformly on $u \in U$, were U is compact in R^q. In many cases, the estimates $\hat{u}_n^\delta\left(\bar{z}_n\right)$ have the same asymptotic properties as well-known M-estimates defined by equations $\hat{u}_n^Q\left(\bar{z}_n\right) = \arg\max_{u \in U} Q_n\left(\bar{z}_n; u\right)$ but often can be much simpler computationally. We consider an algorithmic method for constructing estimates $\hat{u}_n^\delta\left(\bar{z}_n\right)$, which is similar to the accumulation method first proposed by R. Fischer and rigorously developed by L. Le Cam. The main theoretical result of the article is the proof of the theorem, in which conditions of the asymptotic normality of estimates $\hat{u}_n^\delta\left(\bar{z}_n\right)$ are formulated, and the expression is proposed for their matrix of asymptotic mean-square deviations $\lim_{n \to \infty} n \mathrm{E}_{n,u}\left\{ \left(\hat{u}^\delta\left(\bar{z}_n\right) - u\right)\left(\hat{u}^\delta\left(\bar{z}_n\right) - u\right)^T \right\}$.

Keywords: random time series; estimation of distribution parameters; local asymptotical normality; function of estimation quality; asymptotically efficient estimates

1. Introduction. Methods of Construction Asymptotically Efficient Estimates for Parameters of Stationary Time Series

In applications of mathematical statistics to modern problems of data analysis in natural science and technology, it is often impossible to use the classical observation models in the form of a sequence of independent identically distributed random variables (i.i.d. model). As a rule, the i.i.d. model does not provide sufficient accuracy of statistical inferences about the unknown parameters of the investigated physical processes, distorted by noise, if both of them are stationary random processes.

Thus, it is important to generalize the classical results of the statistical theory of parameter estimation, developed for the i.i.d. model, in order to apply them to actual practical problems in the analysis of real physical processes.

In modern systems for analyzing physical wave fields, a large number of parameters are simultaneously measured, and many sensors are used to improve the accuracy of the analysis. That is, *multidimensional* time series $z_t \in R^m$, $t \in \mathbb{Z}$ are subjected to statistical processing, and *vector* parameters are estimated as a result of this processing.

For many statistical models of multivariate time series, it is impossible to synthesize statistically efficient estimates $\hat{u}_n^{\text{ef}}\left(\overline{z}_n\right)$ of vector parameters u for which the standard deviation matrices are minimal for any finite size n of observations and are equal to the inverse Fisher information matrix:

$$K_n^{\text{ef}}(u) = \mathrm{E}_{n,u}\left\{ \left(\hat{u}_n^{\text{ef}}\left(\overline{z}_n\right) - u \right) \left(\hat{u}_n^{\text{ef}}\left(\overline{z}_n\right) - u \right)^{\mathrm{T}} \right\} = J_n^{-1}(u), \qquad (1)$$

where $J_n(u) = \int\limits_{R^{mn}} \left(\nabla_u p_{z,n}\left(\overline{x}_n; u\right) \right) \left(\nabla_u p_{z,n}\left(\overline{x}_n; u\right) \right)^{\mathrm{T}} p_{z,n}^{-1}\left(\overline{x}_n; u\right) d\,\overline{x}_n;$

$$\overline{x}_n = \left(x_1^{\mathrm{T}}, \ldots, x_n^{\mathrm{T}} \right)^{\mathrm{T}} \in R^{mn}; \; \nabla_u p_z\left(\overline{x}_n; u\right) = \left(\frac{\partial}{\partial u_k} p_z\left(\overline{x}_n; u\right), \, k \in \overline{1,q} \right)^{\mathrm{T}};$$

$p_z\left(\overline{x}_n; u\right)$ is the probability density of the observations $\overline{z}_n$.

At the same time, asymptotically efficient (AE) estimates $\hat{u}_n^{\text{ae}}\left(\overline{z}_n\right)$ can be constructed for a wide class of multivariate time series with interdependent elements z_t possessing a strong mixing property [1]. For AE-estimates, equality (1) is attained asymptotically for $n \to \infty$:

$$K^{\text{ae}}(u) = \lim_{n \to \infty} n \mathrm{E}_{n,u}\left\{ \left(\hat{u}_n^{\text{ae}}\left(\overline{z}_n\right) - u \right) \left(\hat{u}_n^{\text{ae}}\left(\overline{z}_n\right) - u \right)^{T} \right\} = \lim_{n \to \infty} n J_n^{-1}(u).$$

They can be found in the class $\mathcal{R}$ of *regular* estimates $\hat{u}\left(\overline{z}_n\right)$ for which the random quantities $\sqrt{n}\left(\hat{u}\left(\overline{z}_n\right) - u \right)$, $u \in U$ have limit distributions with finite second moments. This statement is one of the results of the extensive asymptotic theory of statistical inference for random time series, which is most fully presented in [2]. Fundamental results in this theory were obtained in the known publications [3–6]. In these books, sufficient conditions were established under which AE-estimates exist for many probabilistic models of random time series and continuous processes.

The main condition under which the AE-estimates can be constructed is the local asymptotic normality (LAN) of the likelihood ratio $L_n\left(\overline{z}_n\right)$ of observations $\overline{z}_n$ [3]. It means that the likelihood ratio of the observations $\overline{z}_n$ admits the following asymptotic expansion:

$$L_n\left(\overline{z}_n\right) = ln\frac{p_{z,n}\left(\overline{z}_n; u + n^{-1/2}h\right)}{p_{z,n}\left(\overline{z}_n; u\right)} = h^{\mathrm{T}}\Delta_n\left(\overline{z}_n; u\right) - \frac{1}{2}h^{\mathrm{T}}\Gamma_n(u)h + \alpha_n\left(\overline{z}_n; u, h\right),$$

$$(2)$$

where $\lim_{n \to \infty} \Gamma_n(u) = \Gamma(u) = \lim_{n \to \infty} n^{-1}J_n(u); \Delta_n\left(\overline{z}_n; u\right) \in R^q$ is a family of statistics for which probability distributions tend as $n \to \infty$ to the q-dimensional Gaussian distribu-

tions with the parameters $(0, \Gamma(u))$ uniformly in $u \in U$; $\alpha_n \left(\overline{z}_n ; u, t \right) \to 0$ $(n \to \infty)$ in $dP_n \left(\overline{z}_n \right)$-probability uniformly in $u \in U$; $|h| < c$ where c is any number.

Many publications, for example, [7–14], have been devoted to proving the LAN property for various probabilistic models of time series other than the i.i.d model. The results of research in this direction, obtained up to the end of the twentieth century, are summarized in the monograph [2]. It was shown that the LAN property is inherent in a wide class of multidimensional time series and continuous random processes.

The formulation of the LAN condition (2) largely determined the further development and practical applications of the asymptotic estimation theory. In the well-known monograph [6], it is shown that under the LAN condition, the maximum likelihood estimate belongs to the class $\mathcal{R}$ of regular statistical estimates and is an AE-estimate.

At the same time, using the decomposition (2) of the likelihood function of observations, new AE-estimates were constructed, which differ from the traditional maximum likelihood estimates and are computationally simpler. An elegant and, in many cases, the most computationally simple method for constructing AE-estimates, was proposed in [3,4]. It is based on R. Fisher's [15] idea of "improving" the quality of some "simple" estimate to the quality of an AE-estimate. In mentioned publications, L. Le Cam showed that the AE-estimate can be obtained using the equation:

$$\overset{\wedge \text{ae}}{u_n} \left(\overline{z}_n \right) = u_n^* \left(\overline{z}_n \right) - n^{-1/2} \Gamma_n^{-1} \left(u_n^* \left(\overline{z}_n \right) \right) \Delta_n \left(\overline{z}_n ; u_n^* \left(\overline{z}_n \right) \right), \tag{3}$$

where $u_n^* \left(\overline{z}_n \right)$ is an arbitrary $\sqrt{n}$-consistent estimate of the parameter u for which the quantities $\sqrt{n} \left(u^* \left(\overline{z}_n \right) - u \right)$, $u \in U$, $n \in \mathbb{Z}$ have the property: for any $\varepsilon > 0$ there is $C_\varepsilon > 0$, such that $\sup\limits_{u \in U, n \in \mathbb{Z}^+} \left[P_{n,u} \left\{ \left| \sqrt{n} \left(u^* \left(\overline{z}_n \right) - u \right) \right| > C_\varepsilon \right\} \right] < \varepsilon$.

Note that Equation (3) defines a whole class of AE-estimates, the quality of which is asymptotically equivalent to the quality of the ML-estimate, since $\Delta_n \left(\overline{z}_n ; u \right)$, $\Gamma_n(u)$ in the LAN expansion (2) and the $\sqrt{n}$-consistent estimate $u_n^* \left(\overline{z}_n \right)$ are not unique functions. For this reason, in many practically important cases, formula (3) allows one to obtain AE-estimates, which are computationally much simpler than ML-estimates.

2. Construction of M-Estimates for Parameters of Stationary Time Series with Suitable Asymptotical Properties

The AE-estimates have some disadvantages from the point of view of practical applications. First, they can be synthesized only if the probability density $p_{n,z} \left(\overline{x}_n ; u \right)$ of the observations $\overline{z}_n$ is fully known. In practice, some important details of this density are often not fully defined. Only a certain class $\mathcal{K}$ is known to which this density belongs. Second, the quality of AE-estimates is often unstable to deviations of the actual density $p_{n,z} \left(\overline{x}_n ; u \right)$ from the assumed one for which they were synthesized. Even a small deviation from the expected density can lead to a significant loss in the accuracy of the AE-estimate.

In the publications [16,17], methods were developed for constructing estimates that are *robust* to changes in the distribution of observations, and in many applications, such robust estimates are preferable to AE-estimates. A robust estimate $\hat{u} \left(\overline{z}_n \right)$ is constructed

by finding the global maximum of a certain objective function $Q_n\left(\overline{z}_n; u\right)$ (a criterion of estimation quality), which differs from likelihood function:

$$\hat{u}\left(\overline{z}_n\right) = \underset{u \in U}{arg\ max} Q_n\left(\overline{z}_n; u\right). \tag{4}$$

In addition to robust estimates, estimates synthesized using Equation (4) arise in other problems of mathematical statistics. The examples include Bayesian estimation problems, estimation problems with interfering (nuisance) parameters, problems arising in the analysis of natural and economic dynamical systems.

The estimates obtained as the maxima of some objective functions $Q_n\left(\overline{z}_n; u\right)$ were called "M-estimates". Apart from books [16,17], they were considered in many other publications, for example, in [18,19]. In most of these publications, the M-estimates were constructed and analyzed for the i.i.d. model of random observations.

The authors are not aware of publications in which the asymptotic properties of M-estimates were studied with a sufficient level of mathematical rigor for multidimensional stationary random time series that have a strong mixing property. The authors are also unaware of publications devoted to the construction of computationally simple estimates that are asymptotically equivalent in quality to M-estimates.

In this paper, we consider an approach to solving these problems from the standpoint of view of the asymptotic theory of statistical inference [2], which is based on Le Cam's concept of local asymptotically normality.

We suppose that random objective function $Q_n\left(\overline{z}_n; u\right)$ is twice differentiable in $P_{n,u}$-probability with respect to components of the vector $u \in U$; that is, there exist the following family of vector statistics $d_n\left(\overline{z}_n; u\right)$ and matrix function $F_n\left(\overline{z}_n; u\right)$:

$$
\begin{aligned}
d_n\left(\overline{z}_n; u\right) &= \left(d_{n,k}\left(\overline{z}_n; u\right) = \frac{\partial}{\partial u_k} Q_n\left(\overline{z}_n; u\right), k \in \overline{1, q}\right)^{\mathrm{T}} = \nabla_u Q_n\left(\overline{z}_n; u\right) \in R^q, \\
F_n\left(\overline{z}_n; u\right) &= \left[\frac{\partial}{\partial u_l} d_{n,k}\left(\overline{z}_n; u\right), k, l \in \overline{1, q}\right] = \Delta_u Q_n\left(\overline{z}_n; u\right) \in R^{q \times q}.
\end{aligned}
\tag{5}
$$

In this case, the M-estimate (4) is *one* of the roots $\tilde{u}_n\left(\overline{z}_n\right)$ of the following equation system with respect to the parameter u:

$$d_n\left(\overline{z}_n; u\right) = 0. \tag{6}$$

In this paper, we show how to find the estimate $\hat{u}_n^{\delta}\left(\overline{z}_n\right)$, which is a root of the equation system (6), and, at the same time, it is an $\sqrt{n}$-consistent estimate of the parameter u. It is proved in Theorem 1 that under certain restrictions, such an estimate $\hat{u}_n^{\delta}\left(\overline{z}_n\right)$ can be found using the algorithm

$$\hat{u}_n^{\delta}\left(\overline{z}_n\right) = u_n^*\left(\overline{z}_n\right) - n^{-1/2} \Phi_n\left(u_n^*\left(\overline{z}_n\right)\right) \delta_n\left(\overline{z}_n; u_n^*\left(\overline{z}_n\right)\right), \tag{7}$$

where $\delta_n\left(\overline{z}_n; u\right) = n^{-1/2} d_n\left(\overline{z}_n; u\right)$; $\Phi_n(u) = n^{-1} E_u\left\{F_n\left(\overline{z}_n; u\right)\right\}$; $u_n^*\left(\overline{z}_n\right)$ is any $\sqrt{n}$-consis- tent estimate of the parameter u.

Conditions are formulated in Theorem 1 on the family of statistics $\delta_n\left(\overline{z}_n; u\right)$ and the sequence of the matrix functions $\Phi_n(u)$ that are sufficient for the asymptotic normality of

the estimate (7): $\mathfrak{L}\left\{\sqrt{n}\left(\overset{\wedge\delta}{u}_n\left(\overline{z}_n\right)-u\right)\right\} \to \mathbb{N}(0,D(u))\ (n\to\infty)$, where the asymptotic

covariance matrix $D(u)=\lim\limits_{n\to\infty} n\mathrm{E}_u\left\{\left(\overset{\wedge\delta}{u}_n\left(\overline{z}_n\right)-u\right)\left(\overset{\wedge\delta}{u}_n\left(\overline{z}_n\right)-u\right)^{\mathrm{T}}\right\}$ is equal to

$$D(u)=\Phi^{-1}(u)\Psi(u)\Phi^{-1}(u)\Psi(u)=\lim\limits_{n\to\infty}\mathrm{E}_u\left\{\delta_n\left(\overline{z}_n;u\right)\delta_n^{\mathrm{T}}\left(\overline{z}_n;u\right)\right\}\Phi(u)=\lim\limits_{n\to\infty}\Phi_n(u).$$

The corollary of Theorem 1 describes a method for constructing another estimate $\overset{\sim\delta}{u}_n\left(\overline{z}_n\right)$ that has the same asymptotical distribution as the estimate (7) but does not require an auxiliary $\sqrt{n}$-consistent estimate $u_n^*\left(\overline{z}_n\right)$.

Note that the statements of Theorem 1 and the corollary were formulated earlier in [20]. In our paper, the above statements are proved under more general assumptions, and simpler proofs are given.

Theorem 1. A. *There exists a $\sqrt{n}$-consistent estimate $u_n^*\left(\overline{z}_n\right)$ of the parameter u.*

B. *Let the family of statistics $\delta_n\left(\overline{z}_n,u\right)\in R^m$, $u\in U$, and the sequence of positive definite symmetric $q\times q$-matrix functions $\Phi_n(u)$ satisfy the following constraints:*

B1. *For each value of the parameter $u\in U$, the sequence of statistics $\delta_n\left(\overline{z}_n,u\right)$ is asymptotically normal with zero mean and the covariance matrix $\Psi(u)$:*

$$\mathfrak{L}\left\{\delta_n\left(\overline{z}_n,u\right)\right\}\to\mathcal{N}(0,\Psi(u))(n\to\infty)$$

where $\Psi(u)=\lim\limits_{n\to\infty}\mathrm{E}_u\left\{\delta_n\left(\overline{z}_n,u\right)\delta_n^{\mathrm{T}}\left(\overline{z}_n,u\right)\right\}$.

B2. *For each value of the parameter $u\in U$, the following asymptotic expansion of the statistic $\delta_n\left(\overline{z}_n,u\right)$ holds:*

$$\delta_n\left(\overline{z}_n;u+n^{-1/2}h\right)=\delta_n\left(\overline{z}_n;u\right)+\Phi_n(u)h+\beta_n\left(\overline{z}_n;u,h\right),\ |h|<c\ for\ \forall c;$$

where $\sup\limits_{u\in U,|h|<c}P_{n,u}\left\{\left|\beta_n\left(\overline{z}_n;u,h\right)\right|>\varepsilon\right\}\to 0\ (n\to\infty)$ for any $\varepsilon>0$;

$$\inf\limits_{n\in\mathbb{Z}^+,u\in U}\det\Phi_n(u)>d;\ \lim\limits_{n\to\infty}\sup\limits_{u\in U}\|\Phi_n^{-1}(u)-\Phi^{-1}(u)\|=0;\ \sup\limits_{u\in U}\|\Phi^{-1}(u)\|<C;$$

$\Phi^{-1}(u)$ *is a continuous function of $u\in U$.*
Then the following statement is true:
For any $\sqrt{n}$-consistent estimate $u_n^*\left(\overline{z}_n\right)$ of the parameter $u\in U$, the statistic

$$\overset{\wedge\delta}{u}_n\left(\overline{z}_n\right)=u_n^*\left(\overline{z}_n\right)-n^{-1/2}\Phi_n^{-1}\left(u_n^*\left(\overline{z}_n\right)\right)\delta_n\left(\overline{z}_n;u_n^*\left(\overline{z}_n\right)\right)\tag{8}$$

is the $\sqrt{n}$-consistent and asymptotically normal estimate of the parameter $u\in U$ with the moments $(0,D(u))$:

$$\mathfrak{L}\left\{\sqrt{n}\left(\overset{\wedge\delta}{u}_n\left(\overline{z}_n\right)-u\right)\right\}\to\mathbb{N}(0,D(u))\ (n\to\infty)$$

where $D(u) = \Phi^{-1}(u)\Psi(u)\Phi^{-1}(u)$.

Corollary 1. (a) *Let, for any $n \in \mathbb{Z}^+$, a statistic $\tilde{u}_n^{\delta}\left(\overline{z}_n\right)$ be the root of the equation $\delta_n\left(\overline{z}_n; u\right) = 0$ with respect to the parameter $u \in U$ with probability equal to 1.*

    (b) *Let the statistic $\tilde{u}_n^{\delta}\left(\overline{z}_n\right)$ also is a $\sqrt{n}$-consistent estimate of the parameter $u \in U$. Then the statistic $\tilde{u}_n^{\delta}\left(\overline{z}_n\right)$ is asymptotically normal with the moments $(0, D(u))$.*

Remark 1. (a) *The statement similar to Statement (T1) of Theorem 1 was proved in [3,4] in the case when the objective function $Q_n\left(\overline{z}_n; u\right)$ is the likelihood function of $\overline{z}_n$ having the LAN property (2). In this case $\delta_n\left(\overline{z}_n; u\right) \equiv \Delta_n\left(\overline{z}_n; u\right)$, the matrix function $\Phi_n(u) \equiv \Gamma_n(u)$ and*

$$\mathfrak{L}\left\{\Delta_n\left(\overline{z}_n; u\right)\right\} \to \mathbb{N}(0, \Gamma(u)) \, (n \to \infty); \Gamma(u) = \lim_{n \to \infty} n^{-1}J_n(u),$$

where $J_n(u)$ is the Fisher matrix. It follows from Theorem 1, that in this case

$$D(u) = \Gamma^{-1}(u)\Gamma(u)\Gamma^{-1}(u) = \Gamma^{-1}(u).$$

Consequently, the statistic

$$\hat{u}_n^{\Delta}\left(\overline{z}_n\right) = u_n^*\left(\overline{z}_n\right) - n^{-1/2}\Gamma_n^{-1}\left(u_n^*\left(\overline{z}_n\right)\right)\Delta_n\left(\overline{z}_n; u_n^*\left(\overline{z}_n\right)\right)$$

is asymptotically normal with the parameters $(0, \Gamma(u))$, and hence, it is the asymptotically efficient estimate of the parameter u.

    (b) *It follows from the corollary of Theorem 1 that a statistic $\tilde{u}_n^{\Delta}\left(\overline{z}_n\right)$, which has the property: $\Delta_n\left(\overline{z}_n; \tilde{u}_n^{\Delta}\left(\overline{z}_n\right)\right) = 0$ with probability equal to one, and at the same time is a $\sqrt{n}$-consistent estimate of the parameter $u \in U$, is asymptotically normal with the moments $(0, \Gamma(u))$. Consequently, the statistic $\tilde{u}_n^{\Delta}\left(\overline{z}_n\right)$ is the asymptotically efficient estimate of the parameter $u \in U$.*

Thus, Theorem 1 is, in some sense, an extension of Le Cam's results to the case of an arbitrary objective function $Q_n\left(\overline{z}_n; u\right)$ whose gradient satisfies conditions B1, B2 of Theorem 1.

3. Proof of Theorem 1

In the course of proving Theorem 1, we will omit, if it is obvious, the dependence of functional quantities on the observations $\overline{z}_n$ and sometimes denote their dependence on the parameter u by a subscript.

In these notations, the definition of the estimate $\hat{u}_n\left(\overline{z}_n\right)$ can be written as

$$\hat{u}_n = u_n^* - n^{-1/2}\Phi_n^{-1}(u_n^*)\delta_n(u_n^*).$$

Then we can write the following chain of equalities:

$$
\sqrt{n}\left(\hat{u}_n - u\right) = \sqrt{n}(u_n^* - u) - \Phi_n^{-1}(u_n^*)\delta_n(u_n^*) =
$$
$$
= -\Phi_n^{-1}(u)\delta_n(u) + \left[\sqrt{n}(u_n^* - u) - \Phi_n^{-1}(u_n^*)\delta_n(u_n^*) + \Phi_n^{-1}(u)\delta_n(u)\right] =
$$
$$
= -\Phi_n^{-1}(u)\delta_n(u) + \xi_{n,u}(u_n^*),
$$

(9)

where $\xi_{n,u}(u_n^*) = \sqrt{n}(u_n^* - u) + \Phi_n^{-1}(u_n^*)[-\delta_n(u_n^*) + \delta_n(u)]$. It follows from (9):

$$
\Phi_n(u_n^*)\xi_{n,u}(u_n^*) = -\delta_n(u_n^*) + \delta_n(u) + \Phi_n(u_n^*)\sqrt{n}(u_n^* - u) = \rho_{n,u}(u_n^*).
$$

(10)

By denoting $\tau_{n,u}^* = \sqrt{n}(u_n^* - u)$, we obtain from (10):

$$
\delta_n(u + \tau_{n,u}^*/\sqrt{n}) - \delta_n(u) = \Phi_n(u + \tau_{n,u}^*/\sqrt{n})\tau_{n,u}^* - \rho_{n,u}(\tau_{n,u}^*),
$$

(11)

where the random quantities $\tau_{n,u}^*$, $n \in \mathbb{Z}^+$, $u \in U$ have the property: for any $\varepsilon > 0$ there is $S_\varepsilon > 0$ such that $\sup\limits_{u \in U,\, n \in \mathbb{Z}^+} \left[P_{n,u}\{|\tau_{n,u}^*| > S_\varepsilon\}\right] < \varepsilon$.

At the same time, from condition B2 of Theorem 1, we obtain:

$$
\delta_n\left(\bar{z}_n; u + n^{-1/2}h\right) - \delta_n\left(\bar{z}_n; u\right) = \Phi_n(u)h + \beta_{n,u}\left(\bar{z}_n; h\right),
$$

(12)

where $\sup\limits_{u \in U,\, |h| < c} P_{n,u}\left\{\left|\beta_{n,u}\left(\bar{z}_n; h\right)\right| > \varepsilon\right\} \to 0\ (n \to \infty)$.

The comparison Equations (11) and (12) allow us to prove the following Lemma.

Lemma 1. *Under the conditions of Theorem 1, the following convergences take place for any $\varepsilon > 0$:*
(a) $\lim\limits_{n \to \infty} \sup\limits_{u \in U} P_{n,u}\{|\rho_{n,u}(u_n^*)| > \varepsilon\} = 0$, (b) $\lim\limits_{n \to \infty} \sup\limits_{u \in U} P_{n,u}\{|\xi_{n,u}(u_n^*)| > \varepsilon\} = 0$.

The proof of Lemma 1 is given in Section 5.
The following statement will be needed below.

Lemma 2. *Let some random variables φ_n and η_n have the properties:*
(a) $\lim\limits_{n \to \infty} \mathfrak{L}_n\{\varphi_n\} = \lim\limits_{n \to \infty} P_n\{\varphi_n < x\} = F(x)$; (b) *for any $\varepsilon > 0$* $\lim\limits_{n \to \infty} P_n\{|\eta_n| > \varepsilon\} = 0$.
Then $\lim\limits_{n \to \infty} \mathfrak{L}_n\{\varphi_n + \eta_n\} = \lim\limits_{n \to \infty} P_n\{\varphi_n + \eta_n < x\} = F(x)$.

The proof of Lemma 2 is quite simple, and we omit it.
Taking into account Equations (9)–(12) and statements of Lemmas 1 and 2, we can write the following equalities:

$$
\mathfrak{L}\left\{\sqrt{n}\left(\hat{u}_n - u\right)\right\} = \lim\limits_{n \to \infty} \mathfrak{L}\left\{\Phi_n^{-1}(u)\delta_n(u) + \xi_{u,n}\right\} = \lim\limits_{n \to \infty} \mathfrak{L}\left\{\Phi_n^{-1}(u)\delta_n(u)\right\},
$$

where the existence of the limits follows from conditions B1, B2 of Theorem 1. According to conditions B1 of Theorem 1, we have:
$\lim\limits_{n \to \infty} \mathfrak{L}\{\delta_n(u)\} = \mathbb{N}(0; \Psi(u))$ where $\Psi(u) = \lim\limits_{n \to \infty} E_n\{\delta_n(u)\delta_n^\mathsf{T}(u)\}$
Therefore:
$\lim\limits_{n \to \infty} \mathfrak{L}\{\Phi_n^{-1}(u)\delta_n(u)\} = \mathbb{N}(0; D(u))$, where $D(u) = \Phi^{-1}(u)\Psi(u)\Phi^{-1}(u)$. $\square$

4. Proof of Corollary

Under the conditions B1, B2 of Theorem 1, the statistic $\overset{\wedge\delta}{u}_n\left(\overline{z}_n\right)$ in Equation (8) is asymptotically normal with the moments $(0, \boldsymbol{D}(\boldsymbol{u}))$ for *any* $\sqrt{n}$-consistent estimate $u_n^*\left(\overline{z}_n\right)$. Consequently, due to condition (**b**) of the corollary, the statistic

$$\overset{\wedge\delta}{u}_n\left(\overline{z}_n\right) = \tilde{u}_n^\delta\left(\overline{z}_n\right) + n^{-1/2}\Phi_n^{-1}\left(\tilde{u}_n^\delta\left(\overline{z}_n\right)\right)\delta_n\left(\overline{z}_n; \tilde{u}_n^\delta\left(\overline{z}_n\right)\right)$$

is asymptotically normal with the moments $(0, \boldsymbol{D}(\boldsymbol{u}))$.

But by virtue of condition (**a**) of the corollary, we have that $\overset{\wedge\delta}{u}_n\left(\overline{z}_n\right) = \tilde{u}_n^\delta\left(\overline{z}_n\right)$ with probability equal to one. Hence, the statistic $\tilde{u}_n^\delta\left(\overline{z}_n\right)$ is asymptotically normal with the moments $(0, \boldsymbol{D}(\boldsymbol{u}))$. □

5. Proof of Lemma 1

(**a**) For any $\varepsilon > 0$, $q > 0$ and $\boldsymbol{u} \in U$, we can write the following equation:

$$P_{n,u}\{|\rho_{n,u}(\tau_n^*)| > \varepsilon\} =$$
$$= P_{n,u}\{|\rho_{n,u}(\tau_n^*)| > \varepsilon \cap |\tau_{n,u}^*| \le q\} + P_{n,u}\{|\rho_{n,u}(\tau_n^*)| > \varepsilon \cap |\tau_{n,u}^*| > q\}. \tag{13}$$

Let denote $P_{n,u}\left(\{|\rho_{n,u}(\tau_n^*)| > \varepsilon\}\mid\{|\tau_{n,u}^*| < q\}\right)$ the conditional probability of the event $\{|\rho_{n,u}(\tau_n^*)| > \varepsilon\}$ under the condition of the event $\{|\tau_{n,u}^*| < q\}$. Then (13) can be rewritten as:

$$P_{n,u}\{|\rho_{n,u}(\tau_n^*)| > \varepsilon\} = P_{n,u}\left(\{|\rho_{n,u}(\tau_n^*)| > \varepsilon\}\mid\{|\tau_{n,u}^*| \le q\}\right)P_{n,u}\{|\tau_{n,u}^*| \le q\} +$$
$$+ P_{n,u}\left(\{|\rho_{n,u}(\tau_n^*)| > \varepsilon\}\mid\{|\tau_{n,u}^*| > q\}\right)P_{n,u}\{|\tau_{n,u}^*| > q\}. \tag{14}$$

According to (11), there is $C_\varepsilon > 0$ such that $\sup\limits_{u \in U, n \in \mathbb{Z}^+}\left[P_{n,u}\{|\tau_{n,u}^*| > C_\varepsilon\}\right] < \varepsilon$ for any $\varepsilon > 0$. It follows then from (14) that for any $\varepsilon > 0$ and $\boldsymbol{u} \in U$

$$P_{n,u}\{|\rho_{n,u}(\tau_n^*)| > \varepsilon\} < P_{n,u}\left(\{|\rho_{n,u}(\tau_n^*)| > \varepsilon\}\mid\{|\tau_{n,u}^*| < C_\varepsilon\}\right), \tag{15}$$

where $\rho_{n,u}(\tau_{n,u}^*) = \delta_n(u + \tau_{n,u}^*/\sqrt{n}) - \delta_n(u) - \Phi_n(u + \tau_{n,u}^*/\sqrt{n})\tau_{n,u}^*$.

According to (12), for any $\varepsilon > 0$, $\boldsymbol{u} \in U$ and $|h| < C_\varepsilon$

$$\sup\limits_{u \in U}P_{n,u}\left\{\left|\beta_{n,u}\left(\overline{z}_n; h\right)\right| > \varepsilon\right\} \to 0 \; (n \to \infty), \tag{16}$$

where $\beta_{n,u}\left(\overline{z}_n; h\right) = \delta_n\left(\overline{z}_n; u + n^{-1/2}h\right) - \delta_n\left(\overline{z}_n; u\right) - \Phi_n(u)h$,

It follows from (15), (16) that for any $\varepsilon > 0$ $\lim\limits_{n \to \infty}\sup\limits_{u \in U}P_{n,u}\{|\rho_{n,u}(u_n^*)| > \varepsilon\} = 0$.

(**b**) Since $|\xi_{n,u}(u_n^*)| \le \|\Phi_{n,u}^{-1}(u_n^*)\|\,|\rho_{n,u}(u_n^*)|$, to prove statement (**b**) of Lemma 1, it suffices to check that $\|\Phi_n^{-1}(u_n^*)\|$ is bounded in probability. Since $\boldsymbol{\Phi}_n^{-1}(\boldsymbol{u})$ satisfies conditions B2 of Theorem 1, for any $\varepsilon > 0$ there exists $C_\varepsilon > 0$ that for all n the following inequality holds: $P_{n,u}\{\|\Phi_{n,u}^{-1}(u_n^*)\| \ge C_\varepsilon\} < \varepsilon$. So, we can write:

$$P_{n,u}\{|\xi_{n,u}(u_n^*)| > \varepsilon\} = P_{n,u}\left(\{|\rho_{n,u}(u_n^*)| > \varepsilon\} \cap \left(\|\Phi_{n,u}^{-1}(u_n^*)\| < C_\varepsilon\right)\right) +$$
$$+ P_{n,u}\left(\{|\rho_{n,u}(u_n^*)| > \varepsilon\} \cap \{\|\Phi_{n,u}^{-1}(u_n^*)\| \ge C_\varepsilon\}\right) \le$$
$$\le P_{n,u}\left(\{|\rho_{n,u}(u_n^*)| > \varepsilon\} \cap \left(\|\Phi_{n,u}^{-1}(u_n^*)\| < C_\varepsilon\right)\right) + \varepsilon.$$

Since $|\rho_{u,n}(u_n^*)|$ satisfies statement (**a**) of Lemma 1, one can find a number N_ε such that $\sup\limits_{u \in U, n > N_\varepsilon} P_{u,n}\{|\xi_{u,n}(u_n^*)| > \varepsilon\} < 2\varepsilon$. □

6. Conclusions

The paper investigates the asymptotic properties of statistical estimates for the vector parameter $u \in R^q$ of a stationary multidimensional random time series $z_t \in R^m$, $t \in \mathbb{Z}$ satisfying the strong mixing conditions. We have considered estimates $\tilde{u}\left(\overline{z}_n\right)$ that are solutions of the equations $\nabla_u Q_n\left(\overline{z}_n; u\right) = 0$, $\overline{z}_n = (z_1^{\mathrm{T}}, \ldots, z_n^{\mathrm{T}})^{\mathrm{T}}$, where $Q_n\left(\overline{z}_n; u\right)$ is some objective function for which $\nabla_u Q_n\left(\overline{z}_n; u\right)$ satisfies the constraints of Theorem 1. We have proved that under these constraints, the estimates $\tilde{u}\left(\overline{z}_n\right)$ are $\sqrt{n}$-consistent and asymptotically normal with a limit covariance matrix uniquely determined by the objective function $Q_n\left(\overline{z}_n; u\right)$.

The results of this paper are a generalization of the methods for constructing and analyzing the asymptotic properties of M-estimates, which were previously studied for the case of independent identically distributed observations.

Institutional Review Board Statement: The management of the institute does not object to the publication of the article materials.

Informed Consent Statement: Both authors agreed with content of this paper to be published.

Data Availability Statement: It is theoretical paper, there is no data.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Billingsley, P. Convergence of Probability Measures. In *Wiley Series in Probability and Statistics*; John Wiley and Sons, Inc.: New York, NY, USA, 1999.
2. Taniguchi, M.; Kakizawa, Y. *Asymptotic Theory of Statistical Inferences for Time Series*; Springer Series in Statistics; Springer: New York, NY, USA, 2000.
3. Le Cam, L. *Locally Asymptotically Normal Families of Distributions*; University California Publication Statistics: San Diego, CA, USA, 1960; Volume 3, pp. 37–99.
4. Le Cam, L. *Asymptotic Methods in Statistical Decision Theory*; Springer: New York, NY, USA; Berlin, Germany, 1986.
5. Le Cam, L.; Lo Yang, G. *Asymptotics in Statistics*; Springer: New York, NY, USA, 1990.
6. Ibragimov, I.A.; Has'minskii, R.Z. *Statistical Estimation. Asymptotic Theory. Applications of Mathematics*; Springer: Berlin/Heidelberg, Germany, 1981; Volume 16.
7. Kushnir, A.F. Asymptotically optimal tests for a regression problem of testing hypotheses. *Teor. Veroyatnost. Primenen.* **1968**, *13*, 682–700. (In Russian) [CrossRef]
8. Kushnir, A.F.; Pinskii, A.I. Asymptotically optimal tests of testing hypothesis for an interdependent sample. *Teor. Veroyatnost. Primenen.* **1971**, *16*, 280–291. (In Russian) [CrossRef]
9. Rousas, G.G. *Contiguity of Probability Measures. Some Applications in Statistics*; Cambridge University Press: Cambridge, UK, 1972.
10. Devies, R.B. Asymptotic Inference in Stationary Gaussian Time Series. Advances in applied probability. *Appl. Probab. Trust.* **1973**, *5*, 469–497. [CrossRef]
11. Dzhaparidze, K.O.; Yaglom, A.M. Spectrum parameter estimation in time series analysis. In *Developments in Statistics*; Krishnaiah, P.R., Ed.; Academic Press: New York, NY, USA, 1983; Volume 4, pp. 1–181.
12. Dzhaparidze, K.O. *Parameter Estimations and Hypothesis Testing in Spectral Analysis of Stationary Time Series*; Springer: New York, NY, USA, 1986.
13. Liptser, R.S.; Shiryayev, A.N. Statistics of Random Processes. In *Applications of Mathematics*; Springer: Berlin/Heidelberg, Germany; New York, NY, USA, 1978; Volume 5, p. 6.
14. Kutoyants, Y.A. *Parameter Estimation for Stochastic Processes*; Heldermann: Berlin, Germany, 1984.
15. Fisher, R.A. *Theory of Statistical Estimations, Proceedings of Cambridge Philosophical Society*; Cambridge University Press: Cambridge, UK, 1925; Volume 22, pp. 700–725.
16. Huber, P.J. *Robust Statistics*; Wiley Series in Probability and Statistics; John Wiley and Sons: New York, NY, USA, 1981.
17. Huber, P.J.; Ronchetti, E.M. *Robust Statistics*; Wiley Series in Probability and Statistics; John Wiley and Sons: New York, NY, USA, 2009.
18. Newey, W.K.; McFadden, D. Large Sample Estimation and Hypothesis Testing. In *Handbook of Econometrics*; Engle, R.F., McFadden, D.L., Eds.; Elsevier: Amsterdam, The Netherlands, 1986; Volume 4, Chapter 36.

19. Borovkov, A.A. *Mathematical Statistics*; Gordon and Breach Science Publishers: Amsterdam, The Netherlands, 1998.
20. Kushnir, A.F. Identification algorithms for linear systems with correlated input and output noise. *Probl. Inf. Transm.* **1987**, *23*, 139–150. (In Russian)