ON HADAMARD DIFFERENTIABILITY
AND M-ESTIMATION IN LINEAR MODELS

by

Jian-Jian Ren

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Statistics

Chapel Hill

1990

Approved by:

Advisor

Reader

Reader
JIAN-JIAN REN. On Hadamard Differentiability and M-estimation in Linear Models. (Under the guidance of PRANAB KUMAR SEN)

Abstract: Robust (M-) estimation in linear models generally involves statistical functional processes. For drawing statistical conclusions (in large samples), some (uniform) linear approximations are usually needed for such functionals. In this context, the role of Hadamard differentiability is critically examined in this dissertation. In particular, the concept of the second-order Hadamard differentiability and some related theoretical results are established for the study of the convergence rate in probability of the uniform asymptotic linearity of the M-estimator of regression. Thereby, using Hadamard differentiability through the linear approximation of the estimator, the asymptotic normality, the weak consistency and an asymptotic representation are derived under the weak conditions on the score function $\psi$, the underlying d.f. $F$, and the regression constants.

Some other approaches are also considered for the study of the asymptotic normality and the weak consistency, and these approaches lead to two sets of conditions for the strong consistency of the M-estimators in linear models.
TABLE OF CONTENTS

CHAPTER I. INTRODUCTION
1.1 Introduction ................................................................. 1
1.2 Literature Review .......................................................... 4
1.3 Summary of Results and Further Research ............................. 12

CHAPTER II. ON HADAMARD DIFFERENTIABILITY OF STATISTICAL FUNCTIONAL PROCESSES
2.1 Introduction .................................................................. 19
2.2 Preliminary Notions ......................................................... 24
2.3 Main Results .................................................................. 31
2.4 Proof of Theorem 2.3.1 ..................................................... 35

CHAPTER III. ON M-ESTIMATION IN LINEAR MODELS
3.1 Introduction .................................................................. 63
3.2 Preliminaries .................................................................. 67
3.3 Asymptotic Normality ...................................................... 87
3.4 Consistency .................................................................. 102

CHAPTER IV. ON THE SECOND-ORDER HADAMARD DIFFERENTIABILITY AND ITS APPLICATION TO THE UNIFORM ASYMPTOTIC LINEARITY OF M-ESTIMATION IN LINEAR MODELS
4.1 Introduction .................................................................. 118
4.2 Notations and Assumptions .............................................. 122
4.3 On the Second-Order Hadamard Differentiability ................. 126
4.4 Convergence Rate in Probability of the Uniform Asymptotic Linearity of the M-estimators of Regression .......................... 137

NOTATIONS AND ABBREVIATIONS ........................................ 153

REFERENCES ...................................................................... 155
CHAPTER I

INTRODUCTION

1.1 Introduction

The linear regression model is one of the most widely used tools in statistical analysis. In this dissertation, we consider the following simple linear model:

\[ X_i = \beta^\top c_i + e_i, \quad i = 1, \ldots, n \quad (1.1.1) \]

where the \( c_i \) are known \( p \)-vectors of regression constants, \( \beta = (\beta_1, \ldots, \beta_p)^\top \) is the vector of unknown (regression) parameters, \( p \geq 1 \), and is to be estimated from \( n \) observations \( X_1, \ldots, X_n \), and \( e_i \) are independent and identically distributed random variables (i.i.d.r.v.) with distribution function (d.f.) \( F \). One should notice that (1.1.1) reduces to the classical location model if all \( c_i \), with \( p = 1 \), are equal to 1.

Classically, the problem is solved by minimizing the sum of squares (with respect to \( \theta \)):

\[ \sum_{i=1}^{n} (X_i - c_i^\top \theta)^2 = \min! \quad (1.1.2) \]

or, equivalently, by solving the system of \( p \) equations (with respect to \( \theta \)) obtained
by differentiating (1.1.2)

\[ \sum_{i=1}^{n} c_i (X_i - c_i^T \theta) = 0. \] (1.1.3)

The computation of the Least Square Estimator (LSE) defined by the estimating equations (1.1.3), which are linear equations, can be easily done and the statistical properties of LSE, such as the consistency and the asymptotic normality, were studied in detail (Huber, 1981). But as we know, in spite of its mathematical beauty and computational simplicity, the least square estimator suffers a dramatic lack of robustness against plausible departures from the model based assumptions. In fact, the Influence Function (introduced by Hampel, 1968, 1974), which describes the effect of an additional observation at any point x on a statistics, of the residual of LSE is not bounded. This shows that one single outlier can have an arbitrarily large effect on the estimate (Huber, 1981). Moreover, when the underlying d.f. F has heavy tails (which is a departure from the assumption that the model is usually based on: the normality of e_i), the least square estimator is not efficient, and may not even be consistent.

Due to a lack of robustness of LSE, other robust estimators of \( \beta \) have been proposed. In particular, Huber (1973) robustized the classical equations (1.1.2) and (1.1.3) in a straightforward way: instead of minimizing a sum of squares, we minimize a sum of a less rapidly increasing function \( \rho \) of the residuals, i.e.,

\[ \sum_{i=1}^{n} \rho(X_i - c_i^T \theta) = \min \] (1.1.4)

or, if \( \rho \) is convex with derivative \( \psi \), (1.1.4) is equivalent to
\[ \sum_{i=1}^{n} c_i \psi(X_i - e_i, \theta) = 0. \] (1.1.5)

The resulting estimator of (1.1.5), \( \hat{\beta}_n \), is called \textit{M-estimator of Regression} and \( \psi \) is called \textit{Score Function} of the M-estimator. The robustness properties of such M-estimator were investigated by Hampel, Ronchetti, Rousseeuw and Stahel (1986), and it shows that, for local robustness concern, this M-estimator performs very well. In particular, the influence function of the residual of this M-estimator is bounded, when the score function \( \psi \) is bounded. So, in robustness sense, the M-estimator, comparing with LSE, is an improvement in linear model.

However, the robustness properties of M-estimator of regression is not the interest of our study. Our current concern is the asymptotic properties of this M-estimator: the consistency and the asymptotic normality. As we shall see in Section 1.2, people have required strong conditions either on \( \psi \) or \( F \) for such study. The motive of our research is to study the asymptotic properties of the M-estimator of regression by incorporating Hadamard differentiability under the weak or alternative conditions on \( \psi \), \( F \) and \( \{c_i\}, 1 \leq i \leq n \).

Since M-estimators are implicitly defined by the estimating equations, it is difficult to directly study their statistical properties. Various procedures have been considered by various people for such study. The principle difficulty is the non-linearity of the estimator. Often, such is the case that a linear approximation may provide a very workable tool for the study of the asymptotic properties. In this context, we shall find that Hadamard differentiability provides an easy way to study the asymptotic properties of the M-estimator of regression through the linear approximation of the estimator.
1.2 Literature Review

So far, a considerable amount of work on M-estimator in linear models has been done. Some of the work on the asymptotic properties of the (Huber) M-estimator of regression for fixed $p \geq 1$ will be briefly reviewed in this section.

For convenience sake, denote

$$D_n = \left( c_{ij} \right)_{n \times p}$$ (1.2.1)

as the design matrix, denote

$$\Gamma_n = D_n(D_n^T D_n)^{-1} D_n^T = \left( \gamma_{ij} \right)_{n \times n}$$ (1.2.2)

as the projection matrix, and let

$$\epsilon_n = \max_{1 \leq i \leq n} \gamma_{ii}$$ (1.2.3)

Assume $\hat{\beta}_n = (\hat{\beta}_{n1}, \ldots, \hat{\beta}_{np})^T$ is the M-estimator of regression, i.e. , the solution of (1.1.5). Under the assumptions that the function $\psi$ is continuous, nondecreasing and bounded with a bounded second derivative, Huber (1973) showed that, when $\epsilon p^2 \to 0$,

$$\hat{\alpha}_n = \sum_{j=1}^{p} a_j \hat{\beta}_{nj}$$

is asymptotically normal for all arbitrary coefficients $a_j$ satisfying $\sum_{j=1}^{p} a_j^2 = 1$. He also suggested the estimation of the covariance matrix of $\hat{\beta}_n$. Yohai and Maronna (1979) improved Huber's similar result. They showed that $\hat{\alpha}_n$
is asymptotically normal assuming that $\psi$ has a bounded third derivative and that $\epsilon p^{3/2} \to 0$.

Under weak conditions on $\psi$, Koul (1977) and Jurečková (1977, indirectly) also showed the asymptotic normality, but they both required that the underlying d.f. $F$ has finite Fisher's information, i.e.,

$$0 < I(f) = \int (f'/f)^2 dF < \infty$$

where $F' = f$ is the density function of $F$.

As a corollary of Theorem 4.1 of Bickel (1975), Yohai and Maronna (1979) obtained the asymptotic normality under the following assumptions on $\psi$ and $F$:

(i) $\psi$ is nondecreasing

(ii) There exist positive numbers $b$, $c$, and $d$ such that

$$\frac{\psi(x + z) - \psi(x)}{z} \geq d \quad \text{if } |x| \leq c, \ |z| \leq b$$

and $c$ satisfies $F(c) - F(-c) > 0$.

(iii) $$\int_{-\infty}^{\infty} [\psi(x+h) - \psi(x-h)]^2 dF(x) = o(1), \quad \text{as } h \to 0$$

and for some $\epsilon > 0$

$$\sup_{|q| \leq \epsilon, \ |h| \leq \epsilon} \left\{ \frac{1}{|h|} \int_{-\infty}^{\infty} [\psi(x+q+h) - \psi(x+q)] dF(x) \right\} < \infty.$$ 

(iv) There exists $A(\psi, F)$ such that

$$\int_{-\infty}^{\infty} [\psi(x+h) - \psi(x-h)] dF(x) = h A(\psi, F) + o(1).$$

Relles (1968) proved the consistency of this M-estimator of regression when
\( \psi \) belongs to the Huber family

\[
\psi(x, k) = \min\{|x|, k\} \text{sgn}(x)
\]

where \( k \) is a positive real number. Yohai and Maronna (1979) proved the weak consistency, assuming that, in addition to (i) and (ii), \( \lambda_1(D_n^T D_n) \to \infty \), where \( \lambda_1(A) \) denotes the smallest eigenvalue of the matrix \( A \).

For non-convex \( \rho \), Jurečková (1989) derived the asymptotic normality and the weak consistency under the assumption of certain existence of the second derivative of \( F \).

Among the work on M-estimator, we notice an interesting approach which shall play an important role in our research later on. Treating M-estimator as a statistical functional, its asymptotic properties were considered by Boos and Serfling (1980), Huber (1981) with Fréchet derivative used; by Fernholz (1983, location model) with Hadamard derivative used; and by Kallianpur (1963, MLE) with Gâteaux derivative used. The general idea of these people's work was based on a form of the Taylor expansion involving the derivatives of the functional by Von Mises (1947). In nonparametric models, a parameter \( \theta (=T(F)) \) is regarded as a functional \( T(\cdot) \) on a space \( \mathcal{F} \) of distribution functions \( F \). Thus, the same functional of the sample d.f. \( F_n \) (i.e., \( T(F_n) \)) is regarded as a natural estimator of \( \theta \). Von Mises (1947) expressed \( T(F_n) \) as

\[
T(F_n) = T(F) + T'_F(F_n - F) + \text{Rem}(F_n - F; T(\cdot)) \tag{1.2.4}
\]

where \( T'_F \) is the derivative of the functional \( T \) at \( F \) and \( \text{Rem}(F_n - F; T(\cdot)) \) is the
remainder term in this first order expansion. Note that $F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I(X_i \leq x)$ is based on $n$ independent and identically distributed random variables $X_1, \ldots, X_n$, each having the d.f. $F$, and that $T'_F$ is a linear functional. Hence, $T'_F(F_n - F)$ is an average of $n$ i.i.d.r.v.'s. For drawing statistical conclusions (in large sample), $T'_F$ plays the basic role, and in this context, it remains to show that $\text{Rem}(F_n - F; T(\cdot))$ is asymptotically negligible to the desired extent. Appropriate differentiability conditions are usually incorporated towards this verification. The Fréchet derivative is generally too stringent, and the existence of the second-order derivative is assumed when the Gâteaux derivative is applied (Von Mises, 1947; Kallianpur, 1963) to show that the remainder term is asymptotically negligible, which is a rather strong condition. Since, in a majority of statistical applications, Hadamard differentiability, which is a weaker condition than Fréchet differentiability, suffices without the assumption of the existence of a second-order derivative, we are currently interested in Hadamard differentiability for our research concern.

In particular, Fernholz (1983) considered the M-estimator in location model with Hadamard derivative applied. More specifically, assuming that $X_1, \ldots, X_n$ is a sample from a population with d.f. $F$, Huber (1964) defined the functional corresponding to the estimating equation of the M-estimator of location

$$\sum_{i=1}^{n} \psi(X_i - \theta) = 0 \quad (1.2.5)$$

to be a root $T(F) = \theta$ of the following equation

$$\int \psi(x - \theta) \, dF(x) = 0$$
or equivalently, of

\[ \int_0^1 \psi(F^{-1}(t) - \theta) \, dt = 0 \]  

(1.2.6)

where \( F \) is strictly increasing and continuous. Hence, the M-estimator defined by (1.2.5) is also equivalently defined as a root \( T_n = T(F_n) = \theta \) of

\[ \int_0^1 \psi(F_n^{-1}(t) - \theta) \, dt = 0. \]

We observe that a statistical functional \( T \) induces a functional \( \tau \) on the space \( D[0,1] \) (of right continuous functions having left hand limit) by the relation \( \tau(G) = T(G \circ F) \) for \( G \in D[0,1] \). Thus, (1.2.4) can be written equivalently as

\[ \tau(U_n) = \tau(U) + \tau'(U_n - U) + \text{Rem}(U_n - U; \, \tau(\cdot)), \]  

(1.2.7)

where \( U_n \) is the empirical d.f. of the \( F(X_i) \), \( 1 \leq i \leq n \), and \( U \) is the uniform d.f. on \([0,1] \), and we generally have that, for any \( G \in D[0,1] \), \( \tau(G) = \theta \) is a root of

\[ \int_0^1 \psi(F^{-1}(G^{-1}(t)) - \theta) \, dt = 0 \]  

(1.2.8)

where \( G^{-1} \in D[0,1] \) defined by \( G^{-1}(y) = \inf\{1, \, x; \, G(x) \geq y\} \). Using the Hadamard differentiability (along with some other regularity conditions), Reeds (1976) showed that, for any Hadamard-differentiable (at \( U \)) functional \( \tau \) induced by a statistical functional \( T \),
\[ \sqrt{n} \text{ Rem}(U_n - U; \tau(\cdot)) \xrightarrow{P} 0, \quad \text{as } n \to \infty. \quad (1.2.9) \]

Assuming that \( \psi \) is continuous and piecewise differentiable with a bounded derivative \( \psi' \) which vanishes outside of some bounded interval, Fernholz showed that the functional \( \tau \) defined by (1.2.8) is Hadamard-differentiable at \( U \). Therefore,

\[ \sqrt{n} \left( T(F_n) - T(F) \right) \xrightarrow{P} \sqrt{n} \, \tau'(U_n - U) \xrightarrow{D} N(0, \sigma^2) \]

where \( \tau'(U_n - U) = \sum_{i=1}^{n} IC(X_i; F, T) \) (IC(\( X; F, T \)) is the influence function of \( T \) at \( F \)) and \( \sigma^2 = \text{Var}_F \{ IC(\( X; F, T \)) \} < \infty. \)

We can easily see that Fernholz was dealing with the i.i.d. random variables, but even if we just consider the simple linear regression model with \( p = 1 \)

\[ X_i = t_i \, \theta + e_i, \quad i = 1, \ldots, n \]

\( X_1, \ldots, X_n \) are not i.i.d. random variables. Thus, anything similar to (1.2.8) does not apply to the regression model. Hence, the statistical functional corresponding to the M-estimator of regression cannot be defined in the same fashion as the one for location model.

However, we notice that, for each \( \theta \in \mathbb{R}^p \), the left hand side of the estimating equation (1.1.5), denoted by

\[ E_n(\theta) = \sum_{i=1}^{n} e_i \psi(X_i - \tau_i \theta), \]
is a linear functional of the following empirical function

$$H_n^*(t, \theta) = \sum_{i=1}^{n} c_i I(X_i \leq F^{-1}(t) + c_i \tau \theta),$$

viz.,

$$F_n(\theta) = \int \psi(F^{-1}(t)) \, dH_n^*(t, \theta).$$ \hspace{1cm} (1.2.10)

By the definition of Hadamard differentiability, $F_n(\theta)$ could be the Hadamard derivative of a certain functional $\tau$. Since, as we pointed out earlier, the non-linearity of the estimator is the essential difficulty for the study of the asymptotic properties of the $M$-estimator in linear models, and since the uniform linear approximation of the estimator by Jurečková (1977), i.e., assuming that $F$ has finite Fisher's information, then, for any $K > 0$ and $1 \leq j \leq p$

$$\sup_{\|\theta - \theta^0\| \leq K} |E_j(Y, \theta) - E_j(Y, \theta^0) + \omega(\theta - \theta^0) \tau g^{(j)}(\theta)| \overset{P}{\rightarrow} 0, \hspace{1cm} (1.2.11)$$

where $E_j(Y, \theta)$ are the components of $F_n(\theta)$, $g^{(j)}$ is the $j$th column of $D_n^T D_n$ and $\omega = -\int \psi(x)f'(x) \, dx$, provides an easy access to the study of the asymptotic properties of the $M$-estimator of regression, we may develop a way of using Hadamard differentiability to establish something similar to (1.2.11) for the study of the asymptotic properties of $\hat{\theta}_n$ under the weak conditions on $\psi$, $F$ and $D_n$. More specifically, we may generalize the result (1.2.9) with respect to an empirical function which is similar to $H_n^*(\cdot, \theta)$ (Chapter II), and then, using (1.2.10) for a proper functional $\tau$, we may establish the (Jurečková-) uniform linear
approximation of the M-estimator of regression.

For the more detailed property, Jurečková and Sen (1984) considered the convergence rate in probability of (1.2.11). Assuming that $F$ has finite Fisher's information, that $\psi$ is nondecreasing and continuous with a bounded derivative for $x \in (a,b)$ and $\psi' = 0$ outside of $(a,b)$, and that the design matrix $D_n$ satisfies certain conditions, they showed that, for any $K > 0$,

$$\max_{1 \leq j \leq p} \sup_{\|t\| \leq K} \left| \frac{1}{n} N_{nj}(t) \right| = O_p\left( \frac{1}{\sqrt{n}} \right), \quad (1.2.12)$$

where

$$N_{nj}(t) = \sum_{i=1}^{n} c_{ij} [\psi(X_i - \beta^T e_i + \frac{1}{\sqrt{n}} t^T e_i) - \psi(X_i - \beta^T e_i) - \frac{1}{\sqrt{n}} \gamma^T e_i]$$

and $t \in \mathbb{R}^p$. Furthermore, by (1.2.12), they gave an asymptotic representation of the M-estimator of regression:

$$\sqrt{n}(\hat{\beta}_n - \beta) = Q^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} c_i \psi(X_i - \epsilon_i^T \beta) + O_p\left( \frac{1}{\sqrt{n}} \right)$$

where $c_i^T$ is the row of $D_n$, $\gamma = \int \psi'(x) dF(x)$ and

$$Q = \lim_{n \to \infty} \frac{1}{n} D_n^T D_n$$

is a positive definite $p \times p$ matrix.
1.3 Summary of Results and Further Research

For the simple linear model (1.1.1), we consider a normalized version of the estimating equations (1.1.5) as the following:

\[ M_n(u) = \sum_{i=1}^{n} \epsilon_{ni} \psi(Y_i - \epsilon_{ni}^T u) \]  \hspace{1cm} (1.3.1)

where \( Y_i = X_i - \beta^T c_i \) are i.i.d.r.v.'s with d.f. \( F \), and for every \( n \geq p \),

\[ \epsilon_n = \sum_{i=1}^{n} \epsilon_i \epsilon_i^T = \begin{pmatrix} r_{nij} \end{pmatrix}_{1 \leq i,j \leq p} \] \hspace{1cm} (1.3.2)

\[ \epsilon_n^o = \text{Diag}(\sqrt{r_{nn1}}, \ldots, \sqrt{r_{npp}}) \] \hspace{1cm} (1.3.3)

\[ \epsilon_{ni} = (\epsilon_n^o)^{-1} \epsilon_i = (c_{ni1}, \ldots, c_{nip})^T, \quad 1 \leq i \leq n \] \hspace{1cm} (1.3.4)

\[ u = \epsilon_n^o (\hat{\beta} - \beta) \in \mathbb{R}^p. \] \hspace{1cm} (1.3.5)

Then, we see that (1.1.5) is equivalent to

\[ M_n(u) = 0 \quad \text{(with respect to } u), \] \hspace{1cm} (1.3.6)

i.e., \( \hat{u}_n = \epsilon_n^o (\hat{\beta}_n - \beta) \) is a solution of (1.3.6), and that, for any \( 1 \leq j \leq p \),

\[ \|\epsilon_{nj}\|^2 = 1 \] \hspace{1cm} (1.3.7)

where \( \| \cdot \| \) stands for the Euclidean norm and

\[ \epsilon_{nj} = (c_{n1j}, \ldots, c_{nnj})^T. \] \hspace{1cm} (1.3.8)
Since $M_n(y)$ involves the following empirical function

$$S_n^*(t,y) = \sum_{i=1}^{n} c_{ni} I(Y_i \leq F^{-1}(t) + c_{ni}^T y), \quad t \in [0,1], \ y \in \mathbb{R}^p$$  \hfill (1.3.9)

particularly, for each $y \in \mathbb{R}^p$ and $1 \leq k \leq p$, $M_{nk}(y)$ (the k-th component of $M_n(y)$) is a linear functional of

$$S_{nk}^*(t,y) = \sum_{i=1}^{n} c_{nik} I(Y_i \leq F^{-1}(t) + c_{nik}^T y)$$  \hfill (1.3.10)

(the k-th component of $S_n^*(t,y)$), viz.,

$$M_{nk}(t,y) = \int \psi(F^{-1}(t)) \ dS_{nk}^*(t,y),$$ \hfill (1.3.11)

in Chapter II, we generalized the result (1.2.9) with respect to the following statistical functional process:

$$\left\{ r\left( \frac{S_{nk}^*(\cdot,y)}{\sum_{i=1}^{n} c_{nik}} \right); \ y \in \mathbb{R}^p, \ 1 \leq k \leq p \right\}$$

where $\tau$ is any functional defined on the space $D[0,1]$. More specifically, for any fixed $k=1, \cdots, p$ and any Hadamard differentiable (at $U$) functional $\tau$, we showed that, for any $K>0$, as $n \to \infty$

$$\sup_{|y| \leq K} \left| \sum_{i=1}^{n} c_{nik} \left\{ r\left( \frac{S_{nk}^*(\cdot,y)}{\sum_{i=1}^{n} c_{nik}} \right) - r(U(\cdot)) \right\} - r\left( S_{nk}^*(\cdot,y) - U(\cdot) \sum_{i=1}^{n} c_{nik} \right) \right| \mathbb{P} 0,$$

\hfill (1.3.12)
where $|x| = |(x_1, \ldots, x_q)| = \max_{1 \leq i \leq q} |x_i|$, under the following assumptions:

(A1) $c_{nik} \geq 0, \quad i = 1, 2, \ldots, n$

(A2) $\lim_{n \to \infty} n \max_{1 \leq i \leq n} \| c_{ni} \|^2 < \infty$

(B) $F$ is absolutely continuous with a derivative $F'$ which is positive and continuous with limits at $\pm \infty$.

In the general case of $\{c_{nik}\}$, by (1.3.12), we showed that, for any $1 \leq k \leq p$

$$
\sup_{|y| \leq K} \left| \sum_{i=1}^{n} c^{+}_{nik} r(S^{*+(\cdot, y)} - \sum_{i=1}^{n} c^{-}_{nik} (S^{*-(\cdot, y)} - \sum_{i=1}^{n} c^{-}_{nik} - \tau(U) \sum_{i=1}^{n} c^{+}_{nik} - \tau(U) \sum_{i=1}^{n} \sum_{i=1}^{n} c^{-}_{nik}) \right| \leq P \mathbf{0}, \quad (1.3.13)
$$

under the assumption (B) and

(A3) $\lim_{n \to \infty} n \max_{1 \leq i \leq n} \| \varepsilon^{+}_{ni} \|^2 < \infty, \quad \lim_{n \to \infty} n \max_{1 \leq i \leq n} \| \varepsilon^{-}_{ni} \|^2 < \infty$

where, for every $1 \leq i \leq n, 1 \leq k \leq p, t \in [0,1]$ and $y \in \mathbb{R}^p$,

$$
c_{nik} = c^{+}_{nik} - c^{-}_{nik}; \quad c^{+}_{nik} = \max \{0, c_{nik}\}, \quad c^{-}_{nik} = -\min \{0, c_{nik}\}; \quad (1.3.14)
$$

$$
\varepsilon_{ni} = \varepsilon^{+}_{ni} - \varepsilon^{-}_{ni}; \quad \varepsilon^{+}_{ni} = (c^{+}_{ni1}, \ldots, c^{+}_{nip})^\tau, \quad \varepsilon^{-}_{ni} = (c^{-}_{ni1}, \ldots, c^{-}_{nip})^\tau; \quad (1.3.15)
$$

$$
\varepsilon^{\ast}_{n.k} = \varepsilon^{+}_{n.k} - \varepsilon^{-}_{n.k}; \quad \varepsilon^{+}_{n.k} = (c^{+}_{n1k}, \ldots, c^{+}_{nnk})^\tau, \quad \varepsilon^{-}_{n.k} = (c^{-}_{n1k}, \ldots, c^{-}_{nnk})^\tau; \quad (1.3.16)
$$
\[(d_{nk})^2 = \|\xi_{n,k}^+\|^2, \quad (d_{nk})^2 = \|\xi_{n,k}^-\|^2; \quad (1.3.17)\]

\[\xi_{nik}^+ = \begin{cases} \frac{c_{nik}}{d_{nk}} & \text{if } d_{nk} > 0 \\ 0 & \text{if } d_{nk} = 0, \end{cases} \quad \xi_{nik}^- = \begin{cases} \frac{c_{nik}}{d_{nk}} & \text{if } d_{nk} > 0 \\ 0 & \text{if } d_{nk} = 0; \end{cases} \quad (1.3.18)\]

\[\xi_{ni}^+ = (\xi_{ni1}^+, \ldots, \xi_{nip}^+)^T, \quad \xi_{ni}^- = (\xi_{ni1}^-, \ldots, \xi_{nip}^-)^T; \quad (1.3.19)\]

\[S_{nk}^{++}(t,u) = \sum_{i=1}^{n} c_{nik}^+ I(Y_i \leq F^{-1}(t) + \xi_{ni}^T u); \quad (1.3.20)\]

\[S_{nk}^{--}(t,u) = \sum_{i=1}^{n} c_{nik}^- I(Y_i \leq F^{-1}(t) + \xi_{ni}^T u); \quad (1.3.21)\]

so that

\[S_{nk}^*(t,u) = S_{nk}^{++}(t,u) - S_{nk}^{--}(t,u); \quad (1.3.22)\]

and

\[\|\xi_{n,k}\|^2 = (d_{nk}^+)^2 + (d_{nk}^-)^2 = 1. \quad (1.3.23)\]

Since, for every \(u \in \mathbb{R}^p\) and \(1 \leq k \leq p\), \(S_{nk}^*(\cdot, u)\) is an element of \(D[0,1]\), using (1.3.11), (1.3.13) and the definition of (the first-order) Hadamard differentiability, we, in Chapter III, proved that, for a proper functional \(r\) and any \(K > 0\), as \(n \to \infty\)

\[\sup_{|u| \leq K} \left| \Delta_{\tau_n} + [M_n(u) - M_n(0)] \right| \overset{P}{\to} 0,\]
where $\Delta \tau_n = (\Delta \tau_{n1}, \cdots, \Delta \tau_{np})^\tau$ and

$$
\Delta \tau_{nk} = \sum_{i=1}^{n} c_{nik}^+ \{ \tau \left( \frac{S_{nk}^*(\cdot, y)}{\sum_{i=1}^{n} c_{nik}^+} \right) - \tau \left( \frac{S_{nk}^*(\cdot, 0)}{\sum_{i=1}^{n} c_{nik}^+} \right) \} - \\
- \sum_{i=1}^{n} c_{nik}^- \{ \tau \left( \frac{S_{nk}^*(\cdot, y)}{\sum_{i=1}^{n} c_{nik}^-} \right) - \tau \left( \frac{S_{nk}^*(\cdot, 0)}{\sum_{i=1}^{n} c_{nik}^-} \right) \}, \quad 1 \leq k \leq p.
$$

Furthermore, the uniform asymptotic linearity of M-estimator of regression

$$
\sup_{|u| \leq K} \left| [M_n(u) - M_n(0)] + Q_n u \gamma \right| \overset{P}{\rightarrow} 0 \quad (1.3.24)
$$

was established from showing

$$
\sup_{|u| \leq K} \left| \Delta \tau_n - Q_n u \gamma \right| \overset{P}{\rightarrow} 0,
$$

where $Q_n = \sum_{i=1}^{n} c_{ni} c_n^\top = (C_n^0)^{-1} C_n (C_n^0)^{-1}$ and $\gamma = \int F(x) d\psi(x) > 0$. Immediately, along with some regularity conditions, the asymptotic normality followed from (1.3.24). Using a similar idea, the weak consistency was also shown by incorporating Hadamard differentiability through the uniform asymptotic linearity of the estimator.

However, the Hadamard differentiability approach may not provide a way for the study of the strong consistency, which naturally requires stronger conditions on $\psi$, $F$ and $\{\xi_i\}$. In this context, some other approaches were considered in Chapter III to establish (1.3.24) for the asymptotic normality and the weak consistency, and these approaches led to two sets of conditions for the strong consistency of the M-estimator of regression.
In all of our approaches in Chapter III towards the asymptotic normality and the weak consistency, none of the assumptions, such as bounded second derivative on \( \psi \), or finite Fisher's information on \( F \), or (ii), or any existence of the second derivative of \( F \), was required.

Since, the uniform asymptotic linearity of the M-estimator of regression, (1.3.24), was established by incorporating the first-order Hadamard differentiability, we naturally consider using the second-order Hadamard differentiability, which provides a second order expansion, i.e., a more precise approximation, to study more detailed asymptotic property of (1.3.24): the convergence rate in probability, which leads to an asymptotic representation of the estimator. In Chapter IV, the general definition of the second-order Hadamard differentiability and some related theoretical results, such as the chain rule, certain sufficient conditions of the existence of the second-order Hadamard derivative in some special spaces, and the results corresponding to (1.3.12), (1.3.13) with respect to the second-order Hadamard differentiability, were given. The results on the convergence rate in probability of (1.3.24) by incorporating the second-order Hadamard differentiability and an asymptotic representation of the M-estimator of regression were derived in the last section of Chapter IV, where we basically required weaker conditions on \( \{ \xi_1 \} \), \( \psi \) and \( F \) than Jurečková and Sen (1984).

In further research, we believe that this Hadamard differentiability approach for the study of the asymptotic properties of the M-estimator can be applied to the same study for Random effects models and Mixed effects models. From the fact (Jurečková, 1977) that in linear models, any M-estimate
corresponds to an R-estimate such that the estimates are asymptotically equivalent in the sense of convergence in probability, we also believe that the results of the asymptotic normality and the weak consistency on M-estimator in this dissertation can be established under the same assumptions for the R-estimator by the same Hadamard differentiability approach.
CHAPTER II

ON HADAMARD DIFFERENTIABILITY
OF STATISTICAL FUNCTIONAL PROCESSES

2.1 Introduction

In nonparametric models, a parameter $\theta (=T(F))$ is regarded as a functional $T(\cdot)$ on a space $\mathcal{F}$ of distribution functions (d.f.) $F$. Thus, the same functional of the sample d.f. $F_n$ (i.e., $T(F_n)$) is regarded as a natural estimator of $\theta$. Using a form of the Taylor expansion involving the derivatives of the functional, Von Mises (1947) expressed $T(F_n)$ as

$$T(F_n) = T(F) + T'_F(F_n - F) + \text{Rem}(F_n - F; T(\cdot)) \quad (2.1.1)$$

where $T'_F$ is the derivative of the functional at $F$ and $\text{Rem}(F_n - F; T(\cdot))$ is the remainder term in this first order expansion. Note that $F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I(X_i \leq x)$ is based on $n$ independent and identically distributed random variables (i.i.d.r.v.) $X_1, \ldots, X_n$, each having the d.f. $F$, and that $T'_F$ is a linear functional. Hence, $T'_F(F_n - F)$ is an average of $n$ i.i.d.r.v.'s. For drawing statistical conclusions (in large sample), $T'_F$ plays the basic role, and in this context, it remains to show that $\text{Rem}(F_n - F; T(\cdot))$ is asymptotically negligible to the desired extent.
Appropriate differentiability conditions are usually incorporated towards this verification.

We observe that a statistical functional induces a functional on the space \( D[0,1] \) (of right continuous functions having left hand limits) in the following way:

\[
\tau(G) = T(G \circ F), \quad G \in D[0,1]. \tag{2.1.2}
\]

Thus, (2.1.1) can be written equivalently as

\[
\tau(U_n) = \tau(U) + \tau'_U(U_n - U) + \text{Rem}(U_n - U; \tau(\cdot)), \tag{2.1.3}
\]

where \( U_n \) is the empirical d.f. of the \( F(X_i) \), \( 1 \leq i \leq n \), and \( U \) is the classical uniform d.f. on \([0,1] \) (i.e., \( U(t) = t \), \( 0 \leq t \leq 1 \)). Since the expansion in (2.1.1), written in (2.1.3), is based on some kind of differentiatiion, it is quite natural to inquire about the right form of such a differentiatiion to suit the desired purpose. The current literature is based on an extensive use of the Fréchet derivatives which are generally too stringent. Less restrictive concepts involve the Gâteaux and Hadamard (or compact) derivatives (viz., Kallianpur(1963), Reeds(1976) and Fernholz(1983), among others). Using the Hadamard differentiability (along with some other regularity conditions), Reeds(1976) has shown that

\[
\sqrt{n} \ \text{Rem}(U_n - U; \tau(\cdot)) \overset{\mathcal{P}}{\to} 0, \quad \text{as } n \to \infty \tag{2.1.4}
\]

so that noting that \( \tau'_U(U_n - U) = \frac{1}{n} \sum_{i=1}^{n} IC(X_i; F, T) \), where \( IC(x; F, T) \) is the influence function of \( T \) at \( F \), and assuming that \( \sigma^2 = \text{Var}_F\{IC(X_i; F, T)\} < \infty \), one
obtains that

$$
\sqrt{n} \left( T(F_n) - \theta \right) \overset{\mathbb{P}}{\xrightarrow{\sqrt{n}}} r(U - U) \overset{\mathbb{D}}{\xrightarrow{} } N(0, \sigma^2).
$$

(2.1.5)

In the context of the law of iterated logarithm or some almost sure (a.s.) representation for $T(F_n)$, one may require a stronger mode of convergence in (2.1.4), and this, in turn, may require a more stringent differentiability condition. However, in a majority of statistical applications, Hadamard differentiability suffices, and we shall explore this concept in the context of functional process arising in robust (M-) estimation in simple linear models.

Consider the simple linear model:

$$
X_i = \beta^\top c_i + e_i, \quad i \geq 1
$$

(2.1.6)

where the $c_i$ are known $p$-vectors of regression constants, $\beta = (\beta_1, \ldots, \beta_p)^\top$ is the vector of unknown (regression) parameters, $p \geq 1$, and $e_i$ are i.i.d.r.v.'s with d.f. $F$ ($\in \mathfrak{F}$). Based on a suitable score function $\psi: \mathbb{R} \to \mathbb{R}$, an M-estimator $\hat{\beta}_n$ of $\beta$ is defined as a solution (with respect to $\theta$) of the following equations:

$$
\sum_{i=1}^{n} c_i \psi(X_i - \theta^\top c_i) = 0,
$$

(2.1.7)

where "$\equiv 0$" accommodates the possibility of left hand side being closest to 0 when equality in (2.1.7) is unattainable (such a case may arise when $\psi$ is not continuous everywhere). Setting $Y_i = X_i - \theta^\top c_i$ (i.i.d.r.v.'s with d.f. $F$), we shall see that the following empirical function
\[ S^*_n(t,u) = \sum_{i=1}^{n} \varepsilon_{ni} I(Y_i \leq F^{-1}(t) + \varepsilon_{ni}^T u), \quad t \in [0,1], \; u \in \mathbb{R}^p \]  \hspace{1cm} (2.1.8)

arises typically in the study of the asymptotic properties of \( \hat{\beta}_n \), where the \( \varepsilon_{ni} \) are suitably normalized version of the \( \varepsilon_i \). In particular, we set for every \( n \geq p \),

\[ C_n = \sum_{i=1}^{n} \varepsilon_i \varepsilon_i^T = \begin{pmatrix} r_{nj} \end{pmatrix}_{1 \leq i,j \leq p} \]  \hspace{1cm} (2.1.9)

\[ C_n^0 = \text{Diag}(\sqrt{r_{n11}}, \ldots, \sqrt{r_{np}}) \]  \hspace{1cm} (2.1.10)

\[ \varepsilon_{ni} = (C_n^0)^{-1} \varepsilon_i = (c_{ni1}, \ldots, c_{nip})^T, \quad 1 \leq i \leq n. \]  \hspace{1cm} (2.1.11)

Thus, letting

\[ u = C_n^0 (\ell - \beta) \]  \hspace{1cm} (2.1.12)

and

\[ M_n(u) = \sum_{i=1}^{n} \varepsilon_{ni} \psi(Y_i - \varepsilon_{ni}^T u), \]  \hspace{1cm} (2.1.13)

we see that (2.1.7) is equivalent to

\[ M_n(u) \equiv 0 \quad \text{(with respect to } u). \]  \hspace{1cm} (2.1.14)

The solution of the implicit (set of) equations is greatly facilitated by the following type of (Jurečková-) linearity for M-processes: for every finite real number \( K > 0 \), as \( n \to \infty \).
\[
\sup\{ |M_n(u) - M_n(0) + Q_n(u) \gamma|; \ |u| \leq K \} \overset{P}{\to} 0, \quad (2.1.15)
\]

where \( |y| = (v_1, \ldots, v_q) = \max_{1 \leq i \leq q} |v_i|, \ y \in \mathbb{R}^q \) (q is an arbitrary positive integer),
\( \gamma = \int \psi' dF > 0 \) and \( Q_n = \sum_{i=1}^{n} \varepsilon_{ni} \varepsilon_{ni}^T = (\mathcal{C}_n^0)^{-1} \mathcal{C}_n (\mathcal{C}_n^0)^{-1} \). Under various conditions on the \( \{c_{ni}\} \), the score function \( \psi \) and the d.f. \( F \), (2.1.15) has been established (Jurečková, 1984), and this provides an easy access to the study of the asymptotic properties of the M-estimator \( \hat{\beta}_n \).

We consider a different approach here. Consider a statistical functional process:

\[
\left\{ \frac{S_{nk}^*(\cdot, u)}{\sum_{i=1}^{n} c_{nik}}; \ u \in \mathbb{R}^p, 1 \leq k \leq p \right\}
\]

where \( S_{nk}^*(t, u) = \sum_{i=1}^{n} c_{nik} I(Y_i \leq F^{-1}(t) + \varepsilon_{ni}^T u) \) are the components of \( S_n^*(t, u) \) and \( \tau \) is given by (2.1.2), and for this statistical functional process, the results corresponding to (2.1.4) are shown in this chapter. Since for each \( u \in \mathbb{R}^p \), \( M_{nk}(u) \) (the k-th component of \( M_n(u) \)) is a linear functional of \( S_{nk}^*(t, u) \), viz.,

\[
M_{nk}(t, u) = \int \psi(F^{-1}(t)) \ dS_{nk}^*(t, u), \quad k = 1, \ldots, p, \quad (2.1.16)
\]

\( M_{nk}(u) \) could be the Hadamard derivative of a certain functional \( \tau \). Thus, using the results of this chapter, for a proper functional \( \tau \), we have, for an arbitrary \( K > 0 \), as \( n \to \infty \).
\[
\sup_{|u| \leq K} \left| \Delta \tau_{nk} + \left[ M_{nk}(u) - M_{nk}(0) \right] \right| \overset{P}{\to} 0,
\]  

(2.1.17)

where

\[
\Delta \tau_{nk} = \sum_{i=1}^{n} c_{nik} \left\{ \tau \left( \frac{S_{nk}^{*}(:, u)}{\sum_{i=1}^{n} c_{nik}} \right) - \tau \left( \frac{S_{nk}^{*}(:, 0)}{\sum_{i=1}^{n} c_{nik}} \right) \right\}.
\]  

(2.1.18)

Let

\[
\Delta \tau_n = (\Delta \tau_{n1}, \cdots, \Delta \tau_{np})^T,
\]  

(2.1.19)

then, (2.1.15) follows from showing

\[
\sup_{|u| \leq K} \left| \Delta \tau_n - Q_n u \tau \right| \overset{P}{\to} 0.
\]  

(2.1.20)

Some notations along with basic assumptions are presented in Section 2.2. In the same section, the notion of statistical functionals and the concept of Hadamard differentiability are also introduced. The main results along with part of their derivations are considered in Section 2.3. The proof of Theorem 2.3.1 is given separately in Section 2.4.

### 2.2 Preliminary Notions

Consider the D[0,1] space (of right continuous real valued functions with left hand limits) endowed with the uniform topology. The space C[0,1] of real valued continuous functions, endowed with the uniform topology, is a subspace of
D[0,1]. For every $y \in \mathbb{R}^p$ and $1 \leq k \leq p$, denote by

$$S_{nk}^*(t, y) = \sum_{i=1}^{n} c_{niki} I(Y_i \leq F^{-1}(t) + c_{niki}^T y), \quad t \in [0,1]$$

(2.2.1)

where $Y_i$ are i.i.d. random variables with d.f. $F$ and $c_{niki} \in \mathbb{R}^p$ is given by

$$c_{niki} = (c_{n1i}, \ldots, c_{nip})^T, \quad i=1, \ldots, n.$$  

(2.2.2)

It is easy to see that, for every $y \in \mathbb{R}^p$ and $1 \leq k \leq p$, $S_{nk}^*(\cdot, y)$ is an element of $D[0,1]$. The population counterpart of the $I(Y_i \leq F^{-1}(t) + c_{niki}^T y)$ are the $F(F^{-1}(t) + c_{niki}^T y)$, and this leads us to consider the following:

$$S_{nk}(t, y) = \sum_{i=1}^{n} c_{niki} F(F^{-1}(t) + c_{niki}^T y), \quad 1 \leq k \leq p, \quad t \in [0,1], \quad y \in \mathbb{R}^p.$$  

(2.2.3)

Let $\| \cdot \|$ stands for the Euclidean norm, we also write for every $1 \leq i \leq n$ and $1 \leq k \leq p$,

$$c_{n.k} = (c_{n1k}, \ldots, c_{nkk})^T$$

(2.2.4)

$$c_{niki}^+ = c_{niki}^+ - c_{niki}^-; \quad c_{niki}^+ = \max\{0, c_{niki}\}, \quad c_{niki}^- = \min\{0, c_{niki}\};$$

(2.2.5)

$$c_{n1i}^+ = c_{n1i}^+ - c_{n1i}^-; \quad c_{n1i}^+ = (c_{n1i1}, \ldots, c_{n1ip})^T, \quad c_{n1i}^- = (c_{n1i1}^-, \ldots, c_{n1ip}^-);$$

(2.2.6)

$$c_{n.k}^+ = c_{n.k}^+ - c_{n.k}^- \quad \{ \begin{array}{l}
\end{array} \}$$

(2.2.7)

$$c_{n.k}^+ = (c_{n1k1}, \ldots, c_{nkk})^T, \quad c_{n.k}^- = (c_{n1k1}^-, \ldots, c_{nkk})^T$$


(d^+_nk)^2 = ||c^+_n,k||^2, \quad (d^-_nk)^2 = ||c^-_n,k||^2; \quad (2.2.8)

\varepsilon^+_nik = \begin{cases} 
\frac{c^+_n,k}{d^+_nk} & \text{if } d^+_nk > 0 \\
0 & \text{if } d^+_nk = 0 
\end{cases}, \quad \varepsilon^-_nik = \begin{cases} 
\frac{c^-_n,k}{d^-_nk} & \text{if } d^-_nk > 0 \\
0 & \text{if } d^-_nk = 0 
\end{cases} \quad (2.2.9)

\varepsilon^+_ni = (\varepsilon^+_n1, \ldots, \varepsilon^+_nip)^\tau, \quad \varepsilon^-_ni = (\varepsilon^-_n1, \ldots, \varepsilon^-_nip)^\tau; \quad (2.2.10)

S^{*+}_nk(t,u) = \sum_{i=1}^{n} c^+_n,k I(Y_i \leq F^{-1}(t) + \varepsilon^+_ni,u), \quad t \in [0,1], \quad u \in \mathbb{R}^p \quad (2.2.11)

S^{*-}_nk(t,u) = \sum_{i=1}^{n} c^-_n,k I(Y_i \leq F^{-1}(t) + \varepsilon^-_ni,u), \quad t \in [0,1], \quad u \in \mathbb{R}^p \quad (2.2.12)

so that

S^*_nk(t,u) = S^{*+}_nk(t,u) - S^{*-}_nk(t,u). \quad (2.2.13)

Let

I = (1, \ldots, 1)^\tau, \quad (2.2.14)

and

E_q = [0,1]^q, \quad (2.2.15)

where q is an arbitrary positive integer, then, for any 0 < K < \infty, (t,u) \in E_{p+1},

(t,u,y) \in E_{2p+1}, 1 \leq k \leq p, we write
\[ W_{nk}(t, u) = S^*_{nk}(t, (2u-I)K) - S_{nk}(t, (2u-I)K) \] (2.2.16)

\[ W^{O}_{nk}(t, u) = S_{nk}(t, (2u-I)K) - S_{nk}(t, 0) \] (2.2.17)

\[ G_{nk}(t, u) = \sum_{i=1}^{n} c_{nik} \xi_{ni}^T(2u-I)F'(F^{-1}(t))K \] (2.2.18)

\[ A^*_{nk}(t, u, y) = \sum_{i=1}^{n} c_{nik} I(Y_i \leq F^{-1}(t) + \xi_{ni}^T(2u-I)K - \xi_{ni}^T(2y-I)K) \] (2.2.19)

\[ A_{nk}(t, u, y) = E\{A^*_{nk}(t, u, y)\} \] (2.2.20)

\[ V_{nk}(t, u, y) = A^*_{nk}(t, u, y) - A_{nk}(t, u, y) \] (2.2.21)

\[ V^{O}_{nk}(t, u, y) = A_{nk}(t, u, y) - A_{nk}(t, \frac{1}{2}I, \frac{1}{2}I) \] (2.2.22)

\[ H_{nk}(t, u, y) = \sum_{i=1}^{n} c_{nik} [\xi_{ni}^T(2u-I) - \xi_{ni}^T(2y-I)] F'(F^{-1}(t))K. \] (2.2.23)

It is easy to see that

\[ W_{nk}(t, u) = V_{nk}(t, u, u) = A^*_{nk}(t, u, u) - A_{nk}(t, u, u) \] (2.2.24)

where
\[ A_{nk}^*(t,u,u) = S_{nk}^*(t,(2u-1)K) \]
\[
A_{nk}(t,u,u) = E\{A_{nk}^*(t,u,u)\} = S_{nk}(t,(2u-1)K).
\] (2.2.25)

Also, let \( f \) be a function defined on \( E_q \), we denote

\[
\omega_f(\delta) = \sup_{|\bar{x} - \bar{y}| \leq \delta} |f(\bar{x}) - f(\bar{y})|
\] (2.2.26)

where \( |\bar{x}| = |(x_1, \ldots, x_q)| = \max_{1 \leq i \leq q} |x_i| \) for any positive integer \( q \).

Some assumptions, which may be required for our main results in this chapter, are given below:

(A1) \( c_{nik} \geq 0, \quad i = 1, 2, \ldots, n \)

(A2) \( \|c_{nj}\|^2 = 1, \quad j = 1, \ldots, p; \quad \lim_{n \to \infty} n \max_{1 \leq i \leq n} \|c_{ni}\|^2 < \infty \)

(A3) \( \|c_{nj}\|^2 = 1, \quad 1 \leq j \leq p; \quad \lim_{n \to \infty} n \max_{1 \leq i \leq n} \|c_{ni}^+\|^2 < \infty, \quad \lim_{n \to \infty} n \max_{1 \leq i \leq n} \|c_{ni}^-\|^2 < \infty \)

(B) \( F \) is absolutely continuous with a derivative \( F' \) which is positive and continuous with limits at \( \pm \infty \).

In order to prove our results in Section 2.3, some basic concepts about statistical functional and Hadamard differentiability are needed.
**DEFINITION.** Let $X_1, \ldots, X_n$ be a sample from a population with d.f. $F$ and let $T_n = T_n(X_1, \ldots, X_n)$ be a statistics. If $T_n$ can be written as a functional $T$ of the empirical d.f. $F_n$ corresponding to the sample $X_1, \ldots, X_n$, i.e., $T_n = T(F_n)$, where $T$ does not depend on $n$, then $T$ will be called a Statistical Functional.

For our current study, we actually consider an extended statistical functional, or a statistical functional process:

$$T_{n,y} = T\left(\frac{S^*_{nk}(F(\cdot), y)}{\sum_{i=1}^{n} c_{nik}}\right), \quad y \in \mathbb{R}^p, 1 \leq k \leq p.$$  

The domain of definition of $T$ is assumed to contain $S^*_{nk}(F(\cdot), y)/\sum_{i=1}^{n} c_{nik}$ for all $n \geq 1$, $1 \leq k \leq p$ and $y \in \mathbb{R}^p$, as well as the population d.f. $F$. Usually, the range of $T$ will be the set of real numbers.

As we see earlier, any statistical functional $T$ induces a functional $\tau$ on $D[0,1]$ by the relation given in (2.1.2). In Section 2.3 and 2.4, we will always assume that functional $\tau$ is induced by a statistical functional $T$.

**DEFINITION.** The Influence Function (IF) or Influence Curve (IC) of a statistics $T$ at a d.f. $F$ is usually defined by

$$IC(x; F, T) = \frac{d}{dt} T(F + t(\delta_x - F)) \mid_{t=0}$$

where $\delta_x$ is the d.f. of the point mass one at $x$. 
Let $V$ and $W$ be the topological vector spaces and $L(V, W)$ be the set of continuous linear transformation from $V$ to $W$. Let $\mathcal{A}$ be an open set of $V$,

**DEFINITION.** A functional $T: \mathcal{A} \rightarrow W$ is Hadamard Differentiable (or Compact Differentiable) at $F \in \mathcal{A}$ if there exists $T'_F \in L(V, W)$ such that for any compact set $\Gamma$ of $V$,

$$\lim_{t \to 0} \frac{T(F+tH) - T(F) - T'_F(tH)}{t} = 0$$  \hspace{1cm} (2.2.27)

uniformly for any $H \in \Gamma$. The linear function $T'_F$ is called the Hadamard Derivative of $T$ at $F$.

For convenience sake, in (2.2.27), we usually denote

$$\text{Rem}(tH) = T(F+tH) - T(F) - T'_F(tH),$$  \hspace{1cm} (2.2.28)

then, correspondingly in Section 2.3 and 2.4, we always use the notation

$$\text{Rem}(tH) = \tau(U+tH) - \tau(U) - \tau'_U(tH),$$  \hspace{1cm} (2.2.29)

where $H$ is an arbitrary element of $D[0,1]$.

By the two previous definitions, we can easily see that the existence of Hadamard derivative implies the existence of the influence curve and we also have
\[ T'_F(\delta_x - F) = IC(x; F, T). \] (2.2.30)

2.3 Main Results

**THEOREM 2.3.1.** Suppose \( \tau : D[0, 1] \to \mathbb{R} \) is a functional and is Hadamard differentiable at \( U \). For any fixed \( k = 1, 2, \ldots, p \), assume (A1), (A2) and (B). Then, for any \( K > 0 \), as \( n \to \infty \)

\[
\sup_{|u| \leq K} \left| \sum_{i=1}^{n} c_{n i k} \text{Rem} \left( \frac{S_{n k}^{*}(\cdot, u)}{\sum_{i=1}^{n} c_{n i k}} - U(\cdot) \right) \right| \overset{P}{\to} 0. \quad (2.3.1)
\]

Therefore, we have, as \( n \to \infty \)

\[
\sup_{|u| \leq K} \left| \sum_{i=1}^{n} c_{n i k} \left\{ \tau \left( \frac{S_{n k}^{*}(\cdot, u)}{\sum_{i=1}^{n} c_{n i k}} \right) - \tau(U(\cdot)) \right\} - \tau'(U(\cdot)) \sum_{i=1}^{n} c_{n i k} \right| \overset{P}{\to} 0. \quad (2.3.2)
\]

The proof of Theorem 2.3.1 will be given in Section 2.4. In the general case of \( \{c_{n i k}\} \), we have the following theorem.

**THEOREM 2.3.2.** Suppose \( \tau : D[0, 1] \to \mathbb{R} \) is a functional and is Hadamard differentiable at \( U \). Assume (A3) and (B). Then, for any \( 1 \leq k \leq p \) and \( K > 0 \), as \( n \to \infty \).
\[
\sup_{|u| \leq K} \left| \sum_{i=1}^{n} c_{nik}^{*} \frac{S_{n}^{*}(\cdot, y)}{\sum_{i=1}^{n} c_{nik}^{*}} - \sum_{i=1}^{n} c_{nik}^{*} \frac{S_{n}^{*}(\cdot, y)}{\sum_{i=1}^{n} c_{nik}^{*}} \right| - \tau(U) \sum_{i=1}^{n} c_{nik}^{*} - \tau(U) \sum_{i=1}^{n} c_{nik}^{*} - U(\cdot) \sum_{i=1}^{n} c_{nik}^{*} \right| \leq 0. \tag{2.3.3}
\]

Proof. Consider any \(1 \leq k \leq p\). If \(d_{nk}^{+} = 0\) or \(d_{nk}^{-} = 0\), then this simply is the case of Theorem 2.3.1, and (2.3.3) is just the same as (2.3.2) since

\[
\sum_{i=1}^{n} c_{nik}^{*} \frac{S_{n}^{*}(\cdot, y)}{\sum_{i=1}^{n} c_{nik}^{*}} \quad \text{or} \quad \sum_{i=1}^{n} c_{nik}^{*} \frac{S_{n}^{*}(\cdot, y)}{\sum_{i=1}^{n} c_{nik}^{*}}
\]

is equal to 0. If \(d_{nk}^{+} > 0\) and \(d_{nk}^{-} > 0\), let

\[
\xi_{ni}^{\tau} = (\xi_{ni}^{*}, -\xi_{ni}^{-}) \in \mathbb{R}^{2p}. \tag{2.3.4}
\]

By the assumption (A3), we have

\[
\lim_{n \to \infty} n \max_{1 \leq i \leq n} \|\xi_{ni}\|^2 < \infty.
\]

Note that

\[
\sum_{i=1}^{n} (\xi_{nik}^{*})^2 = 1, \quad \sum_{i=1}^{n} (\xi_{nik}^{-})^2 = 1,
\]

and

\[
\sum_{i=1}^{n} c_{nik}^{2} = (d_{nk}^{+})^2 + (d_{nk}^{-})^2 = 1.
\]
therefore, for both \( \{ \xi_{ni}^+, \xi_{ni}^- \} \) and \( \{ \xi_{ni}^+, \xi_{ni}^- \} \), (A1) and (A2) are satisfied, and we have

\[
0 < d_{nk}^+, d_{nk}^- < 1.
\]

We also observe, for \(|u| \leq K\),

\[
\frac{S_{nk}^{*}(t, u)}{\sum_{i=1}^{n} c_{nik}^+} = \frac{\sum_{i=1}^{n} c_{nik}^+ I(Y_i \leq F^{-1}(t) + c_{ni}^+ u_1)}{\sum_{i=1}^{n} c_{nik}^+} = \frac{\sum_{i=1}^{n} c_{nik}^+ I(Y_i \leq F^{-1}(t) + c_{ni}^+ u_1 - c_{ni}^- u_2)}{\sum_{i=1}^{n} c_{nik}^+}, \quad |u_1|, |u_2| \leq K
\]

\[
= \sum_{i=1}^{n} c_{nik}^+ I(Y_i \leq F^{-1}(t) + c_{ni}^+ u_1 - c_{ni}^- u_2), \quad \forall \in \mathbb{R}^{2p}, \quad |\forall| \leq K.
\]

Therefore, by Theorem 2.3.1, we have, for any \( K > 0 \), as \( n \to \infty \)

\[
\sup_{|u| \leq K} \left| \sum_{i=1}^{n} c_{nik}^+ \left\{ \tau\left( \frac{S_{nk}^{*}(\cdot, u)}{\sum_{i=1}^{n} c_{nik}^+} \right) - \tau(U(\cdot)) \right\} - \tau U\left( \frac{S_{nk}^{*}(\cdot, u)}{d_{nk}^+} - U(\cdot) \sum_{i=1}^{n} c_{nik}^+ \right) \right| \to 0.
\]

Since \( 0 < d_{nk}^+ < 1 \), we have, as \( n \to \infty \)

\[
\sup_{|u| \leq K} \left| \sum_{i=1}^{n} c_{nik}^+ \left\{ \tau\left( \frac{S_{nk}^{*}(\cdot, u)}{\sum_{i=1}^{n} c_{nik}^+} \right) - \tau(U(\cdot)) \right\} - \tau U\left( S_{nk}^{*}(\cdot, u) - U(\cdot) \sum_{i=1}^{n} c_{nik}^+ \right) \right| \to 0.
\]

(2.3.5)
Similarly, we can show that, as $n \to \infty$,

$$
\sup_{|u| \leq K} \left| \sum_{i=1}^{n} c_{nik} \{ r \left( \frac{S_{nk}^{*} (\cdot, u)}{\sum_{i=1}^{n} c_{nik}} \right) - r(U(\cdot)) \} - r \left( \frac{S_{nk}^{*} (\cdot, u)}{\sum_{i=1}^{n} c_{nik}} \right) - U(\cdot) \sum_{i=1}^{n} c_{nik} \right| \to 0.
$$

(2.3.6)

Therefore, (2.3.3) follows from (2.3.5) and (2.3.6).

COROLLARY 2.3.3. Suppose $\tau : D[0, 1] \to \mathbb{R}$ is a functional and is Hadamard differentiable at $U$. Assume (A2) and (B), and assume

$$
\min_{1 \leq j \leq p} \{ \lim_{n \to \infty} d_{nj}^{+} \} > 0, \quad \min_{1 \leq j \leq p} \{ \lim_{n \to \infty} d_{nj}^{-} \} > 0.
$$

Then, for any $1 \leq k \leq p$ and $K > 0$, we have (2.3.3).

Proof. By the assumptions on $d_{nj}^{+}$ and $d_{nj}^{-}$, there exists a real number $c > 0$ such that $d_{nj}^{+} > c$ and $d_{nj}^{-} > c$, for all $n \geq 1$ and $1 \leq j \leq p$. By (A2), we have

$$
\lim_{n \to \infty} n \max_{1 \leq i \leq n} \| e_{ni}^{+} \|^2 \leq \frac{1}{c^2} \lim_{n \to \infty} n \max_{1 \leq i \leq n} \| e_{ni}^{-} \|^2 < \infty,
$$

and

$$
\lim_{n \to \infty} n \max_{1 \leq i \leq n} \| e_{ni}^{-} \|^2 \leq \frac{1}{c^2} \lim_{n \to \infty} n \max_{1 \leq i \leq n} \| e_{ni}^{-} \|^2 < \infty.
$$

Hence, (A3) holds. Therefore, (2.3.3) follows from Theorem 2.3.2. \qed
2.4 Proof of Theorem 2.3.1

In this section, for any fixed \(k=1, \ldots, p\), (A1), (A2) and (B) are assumed. First, we notice that, for a fixed \(k=1, \ldots, p\), (A1) and (A2) imply three facts: there exists \(M>0\) such that for \(n\) large enough and any \(1 \leq j \leq p\),

\[
\left(\sum_{i=1}^{n} |c_{nij}|^{2}\right) \max_{1 \leq j \leq p} \max_{1 \leq i \leq n} |c_{nij}| \leq M, \tag{2.4.1}
\]

\[
\max_{1 \leq i \leq n} c_{nij}^{2} \rightarrow 0, \quad \text{as } n \rightarrow \infty, \tag{2.4.2}
\]

and further

\[
\sum_{i=1}^{n} c_{nik} \rightarrow \infty, \quad \text{as } n \rightarrow \infty. \tag{2.4.3}
\]

We also notice that (B) implies that \(F'\) is bounded and uniformly continuous. The result (2.3.1) will first be proved on \(C[0,1]\), and then be extended to \(D[0,1]\).

For any \(0<K<\infty\), consider

\[
S_{nk}^{*}(t,y) = \sum_{i=1}^{n} c_{nik} I(Y_{i} \leq F^{-1}(t)+\xi_{n}^{T}u), \quad t \in [0,1], |u| \leq K.
\]

For each \(i\):

\[
I(Y_{i} \leq F^{-1}(t)+\xi_{n}^{T}u) = \begin{cases} 
1 & \text{if } F(Y_{i}-\xi_{n}^{T}u) \leq t \\
0 & \text{otherwise},
\end{cases}
\]
the curve \( l_{ni} : t = F(Y_i - c_{ni}^T \bar{u}) \) is continuous in \( \bar{u} \), because \( F \) is continuous. Hence, for each \( n \geq 1 \) and \( Y_1, \ldots, Y_n \), \([0, 1] \times [-K, K]^P \) is divided into finite pieces by smooth curves \( l_{ni} \), \( 1 \leq i \leq n \), and the value of \( S_{nk}^*(t, \bar{u}) \) is simply a constant in each different piece, or region. Also, if at most \( r \) curves \( l_{nij} \), \( j = 1, \ldots, r \), intersect at one point, the largest jump of \( S_{nk}^*(t, \bar{u}) \) is \( \sum_{j=1}^{r} \epsilon_{nij} \). Let \( \tilde{S}_{nk}^*(t, \bar{u}) \) (got by smoothing \( S_{nk}^*(t, \bar{u}) \) through the above regions) be a continuous version (in \( (t, \bar{u}) \)) of \( S_{nk}^*(t, \bar{u}) \), then \( \tilde{S}_{nk}^*(t, (2 \bar{u} - I)K) \) is an element of \( C[0, 1]^{P+1} \), and we have the following lemma.

**Lemma 2.4.1.** For any fixed \( k = 1, \ldots, p \), assume (A1). Then,

\[
\sup_{(t, \bar{u}) \in E_1 \times [-K, K]^P} \left| \tilde{S}_{nk}^*(t, \bar{u}) - S_{nk}^*(t, \bar{u}) \right| \leq (p+1) \max_{1 \leq i \leq n} \epsilon_{nik}, \quad \text{a.s.}
\]

(2.4.4)

**Proof.** Without loss of the generality, suppose that \( l_{ni} \), \( 1 \leq i \leq p+2 \), intersect at one point \((t_0, \bar{u}_0)\). Since \( F \) is strictly increasing, we have

\[
Y_1 - \epsilon_{n1}^T \bar{u}_0 = Y_2 - \epsilon_{n2}^T \bar{u}_0 = \cdots = Y_{p+2} - \epsilon_{n(p+2)}^T \bar{u}_0
\]

(2.4.5)

which are \((p+1)\) linear equations with respect to \( \bar{u}_0 \in \mathbb{R}^p \). Then, there exist two \( p \times p \) matrices \( A_1 \) and \( A_2 \), depending on \( \epsilon_{ni} \), \( 1 \leq i \leq (p+2) \), such that

\[
\bar{u}_0 = A_1 (Y_1, \ldots, Y_p)^T - A_2 (Y_2, \ldots, Y_{p+1})^T
\]
\[ = A_2(Y_2, \ldots, Y_{p+1})^T - A_2(Y_3, \ldots, Y_{p+2})^T. \quad (2.4.6) \]

Since (2.4.5) and (2.4.6) imply that \( Y_1, \ldots, Y_{p+2} \) are linearly dependent, i.e.,

\[ Y_1 - Y_2 = (\xi_n1 - \xi_n2)^T A_2(Y_2 - Y_3, \ldots, Y_{p+1} - Y_{p+2})^T, \]

and since \( F \) is continuous, then

\[ P\{ Y_1 - Y_2 = (\xi_n1 - \xi_n2)^T A_2(Y_2 - Y_3, \ldots, Y_{p+1} - Y_{p+2})^T \} = 0. \]

For each \( n \), the probability that more than \((p+1)\) \( F \)'s intersect at one point only depend on \( \{ c_{nij} \} \). Hence, with probability one, no more than \((p+1)\) \( F \)'s intersect at one point for each \( n \geq 1 \). Therefore, with probability one, the largest jump of \( S^*_{nk}(t, u) \) is no larger than \((p+1)\max_{1 \leq i \leq n} c_{nik}\). Since \( \tilde{S}^*_{nk}(t, u) \) is a continuous smoothing of \( S^*_{nk}(t, u) \), (2.4.4) follows. \( \square \)

**Lemma 2.4.2.** Assume (A2) and (B), then, for any \( 1 \leq k \leq p \), as \( n \to \infty \)

\[ \sup_{(t, u, y) \in E_{2p+1}} \left| V^\circ_{nk}(t, u, y) - H_{nk}(t, u, y) \right| \to 0, \quad (2.4.7) \]

and

\[ \omega_{H_{nk}}(\delta) \to 0, \quad \text{as } \delta \to 0 \quad (2.4.8) \]
uniformly with respect to \( n \geq 1 \) and \( 1 \leq k \leq p \).

Proof. For any \( (t, u, y) \in E_{2p+1} \), we have

\[
V_{nk}^O(t, u, y) - H_{nk}(t, u, y)
= \sum_{i=1}^{n} c_{niki}(F'(\xi) - F'(F^{-1}(t)))\left[c_{ni1}^+ \tau(2u - I) - c_{ni1}^- \tau(2y - I)\right]K
\]

where \( \xi \) is between \( F^{-1}(t) \) and \( (F^{-1}(t) + [c_{ni1}^+ \tau(2u - I) - c_{ni1}^- \tau(2y - I)]K) \). Hence,

\[
\sup_{(t, u, y) \in E_{2p+1}} |V_{nk}^O(t, u, y) - H_{nk}(t, u, y)|
\leq K \sum_{i=1}^{n} |c_{niki}| \left|\begin{array}{c} c_{ni1}^+ \ldots + |c_{nipl}| \end{array}\right| \sup |F'(\xi) - F'(F^{-1}(t))|
\leq K \sum_{i=1}^{n} \sum_{j=1}^{p} |c_{niki}| |c_{niij}| \sup |F'(\xi) - F'(F^{-1}(t))|
\leq K \sum_{j=1}^{p} \sqrt{\sum_{i=1}^{n} c_{niki}^2} \sqrt{\sum_{i=1}^{n} c_{niij}^2} \sup |F'(\xi) - F'(F^{-1}(t))|
= pK \sup |F'(\xi) - F'(F^{-1}(t))|.
\]

Then, (2.4.7) follows from (2.4.2) and the uniform continuity of \( F' \).

For any \( (t, u, y), (s, x, y) \in E_{2p+1} \), we have

\[
H_{nk}(t, u, y) - H_{nk}(s, x, y)
\]
\[ = \sum_{i=1}^{n} c_{n1}^{+} c_{ni}^{+} T(2y - 1) - c_{ni}^{+} T(2y - 1) \right] \{ F'(F^{-1}(t)) - F'(F^{-1}(s)) \} K + \]
\[ + \sum_{i=1}^{n} c_{n1}^{+} c_{ni}^{+} T(y - x) - c_{ni}^{+} T(y - y) \} F'(F^{-1}(s)) 2K. \]

By the similar proof above,

\[ \omega_{nk}(\delta) \leq K \sum_{i=1}^{n} |c_{n1}^{+}|\left(|c_{ni1}| + \ldots + |c_{nip}|\right) \sup_{|t-s| \leq \delta} |F'(F^{-1}(t)) - F'(F^{-1}(s))| + \]
\[ + 2K \delta M_1 \sum_{i=1}^{n} |c_{n1}^{+}| \left(|c_{ni1}| + \ldots + |c_{nip}|\right) \]
\[ \leq pK \sup_{|t-s| \leq \delta} |F'(F^{-1}(t)) - F'(F^{-1}(s))| + 2pKM_1 \delta \]

where \(M_1\) is a bound of \(F'\). Therefore, (2.4.8) follows from the uniform continuity of \(F'\).

\[ \square \]

**COROLLARY 2.4.3.** Assume (A2) and (B). Then, for any \(1 \leq k \leq p\), we have, as \(n \to \infty\)

\[ \sup_{(t,u) \in E_{p+1}} \left| W_{nk}^0(t,u) - G_{nk}(t,u) \right| \to 0, \quad (2.4.9) \]

and

\[ \omega_{nk}(\delta) \to 0, \quad \text{as} \ \delta \to 0. \quad (2.4.10) \]
Proof. For any \((t, y) \in E_{p+1}\), we have

\[ W_{nk}(t, y) = V_{nk}(t, y, y), \quad G_{nk}(t, y) = H_{nk}(t, y, y). \]

Therefore, (2.4.9) and (2.4.10) follow from Lemma 2.4.2. \(\Box\)

Let

\[ L_q(r) = \{ l_1, \ldots, l_q \} \mid l_i = 0, 1, \ldots, r; i = 1, 2, \ldots, q \]

where \( r \) is an arbitrary positive integer.

**Lemma 2.4.4. (Neuhaus)** Let \( \delta > 0 \), \( m(\delta) = \max\{ r | \delta 2^r \leq 1 \} \), \( m > m(\delta) \) and a function \( f : E_q \to \mathbb{R} \) be given. Then, for points \( t_1, t_2 \in L_q(2^m) \) with \( |t_1 - t_2| \leq 2^{-m(\delta)} \), the following inequality holds

\[ |f(t_1) - f(t_2)| \leq 4q \sum_{r=m(\delta)}^{2^m} \sum_{\mu=1}^{q} \sup_{j} |f(j) - f(j + e_{\mu}2^{-r})| \]

where the 'sup' is to be taken over all \( j \in L_q(2^r) \) with \( j + e_{\mu}2^{-r} \in E_q \) (\( e_{\mu} \) denotes the \( \mu \)-th unit-vector of \( E_q \)).

For \( x = (x_1, \ldots, x_q) \), \( y = (y_1, \ldots, y_q) \in E_q \), we denote
\[ x \preceq y \quad \text{iff} \quad x_i \leq y_i, \quad \text{for } i = 1, 2, \ldots, q \quad (2.4.11) \]

**Proposition 2.4.5.** For any fixed \( k = 1, 2, \ldots, p \), assume (A1), (A2) and (B). Then, for any \( \epsilon > 0 \),

\[
\lim_{\delta \to 0} \lim_{n \to \infty} P(\omega_{W_{nk}}(\delta) \geq \epsilon) = 0. \quad (2.4.12)
\]

**Proof.** For any \( (t, y), (s, x) \in E_{p+1} \) with \( |(t, y) - (s, x)| < \delta \), there exist \( (t', y'), (t'', y''), (s', x'), (s'', x'') \in L_{p+1}(2^m) \), where \( m \) is an arbitrary positive integer, such that

\[
(t', y') \preceq (t, y) \preceq (t'', y''); \quad (s', x') \preceq (s, x) \preceq (s'', x'')
\]

and

\[
|(t', y') - (t, y)| \leq 2^{-m}, \quad |(t'', y'') - (t, y)| \leq 2^{-m};
\]

\[
|(s', x') - (s, x)| \leq 2^{-m}, \quad |(s'', x'') - (s, x)| \leq 2^{-m}.
\]

We write

\[
W_{nk}(t, y) - W_{nk}(s, x) = [W_{nk}(t, y) - W_{nk}(t', y')] +
\]

\[ + [W_{nk}(t', y') - W_{nk}(s'', y'')] + [W_{nk}(s'', y'') - W_{nk}(s, x)] \]
and notice that

\[ |(t', y') - (s'', y'')| \leq |(t', y') - (t, y)| + |(t, y) - (s, y)| + |(s, y) - (s'', y'')| \leq 2^{-m+1}. \]

Consider \([W_{nk}(t, y) - W_{nk}(t', y')].\) Since \(A_{nk}^*(t, y, y)\) is monotonic in each component of \((t, y, y),\)

\[ [W_{nk}(t, y) - W_{nk}(t', y')] \]

\[ = [A_{nk}^*(t, y, y) - A_{nk}^*(t', y', y')] - [A_{nk}(t, y, y) - A_{nk}(t', y', y')] \]

\[ \leq [A_{nk}^*(t'', y'', y') - A_{nk}^*(t', y', y')] - [A_{nk}(t, y, y) - A_{nk}(t', y', y')] \]

\[ = [V_{nk}(t'', y'', y') - V_{nk}(t', y', y')] - [A_{nk}(t, y, y) - A_{nk}(t', y', y')] \]

\[ = [V_{nk}(t'', y'', y') - V_{nk}(t', y', y')] - [V_{nk}^o(t, y, y) - V_{nk}^o(t'', y'', y')] - \]

\[ - (t - t'') \sum_{i=1}^{n} c_{nik} \]

and

\[ [W_{nk}(t, y) - W_{nk}(t', y')] \]

\[ \geq [A_{nk}^*(t', y', y'') - A_{nk}^*(t', y', y')] - [A_{nk}(t, y, y) - A_{nk}(t', y', y')] \]
\[ \begin{align*}
&= [V_{nk}(t', y', y'') - V_{nk}(t', y', y')] - [A_{nk}(t, y, y) - A_{nk}(t', y', y'')] \\
&= [V_{nk}(t', y', y'') - V_{nk}(t', y', y')] - [V_{nk}^o(t, y, y) - V_{nk}^o(t', y', y'')] - \\
&\quad - (t - t') \sum_{i=1}^{n} c_{nik}
\end{align*} \]

hence,

\[ |W_{nk}(t, y) - W_{nk}(t', y')| \leq 2^m \sum_{i=1}^{n} c_{nik} + \]

\[ + \sup \{ |V_{nk}(t, y, y) - V_{nk}(s, \bar{x}, \bar{y})|; \]

\[ (t, y, y), (s, \bar{x}, \bar{y}) \in L_{2p+1}(2^m), |(t, y, y) - (s, \bar{x}, \bar{y})| \leq \delta + 2^{-m+1} \} \]

\[ + \sup \{ |V_{nk}^o(t, y, y) - V_{nk}^o(s, \bar{x}, \bar{y})|; |(t, y, y) - (s, \bar{x}, \bar{y})| \leq 2^m \}. \]

We can treat \([W_{nk}(s'', y'') - W_{nk}(s, y)]\) in the similar way. Therefore, we have

\[ \omega_{W_{nk}}(\delta) \leq 2 \omega_{V_{nk}}^o(2^m) + 2^m \sum_{i=1}^{n} c_{nik} + \]

\[ + 3 \sup \{ |V_{nk}(t, y, y) - V_{nk}(s, \bar{x}, \bar{y})|; \]

\[ (t, y, y), (s, \bar{x}, \bar{y}) \in L_{2p+1}(2^m), |(t, y, y) - (s, \bar{x}, \bar{y})| \leq \delta + 2^{-m+1} \}. \quad (2.4.13) \]
Since,
\[
\omega^0_{nk}(2^{-m}) \leq 2 \sup_{(t,υ,γ) ∈ E_{2p+1}} |V^0_{nk}(t,υ,γ) - H_{nk}(t,υ,γ)| + \omega_{nk}(2^{-m}),
\]
by Lemma 2.4.2, we have
\[
\omega^0_{nk}(2^{-mn}) \to 0, \quad \text{as } n \to \infty \quad (2.4.14)
\]
where
\[
m_n = \max\{m; \frac{1}{a_n} \geq 2^{cm}\}, \quad (2.4.15)
\]
\[
c = (4p+3)/(4p+4) \quad \text{and} \quad a_n = \max_{1 \leq j \leq p} \max_{1 \leq i \leq n} |c_{nij}| \quad (2.4.16)
\]
so that \(m_n \to \infty\), as \(n \to \infty\). By (2.4.1),
\[
2^{-mn} \sum_{i=1}^{m} c_{nik} \leq M (2^{mn} a_n)^{-1} \leq M 2^{-mn} 2^{c(mn+1)}
\]
\[
= M 2^{-\frac{mn}{4(p+1)} + c} \to 0, \quad \text{as } n \to \infty. \quad (2.4.17)
\]
Hence, by virtue of (2.4.13), (2.4.14) and (2.4.17), it suffices to show that
\[
\lim_{δ \to 0} \lim_{n \to \infty} P(\sup\{|V_{nk}(t,υ,γ) - V_{nk}(s,χ,γ)|; (t,υ,γ), (s,χ,γ) ∈ L_{2p+1}(2^{-mn}), |(t,υ,γ) - (s,χ,γ)| ≤ δ\} ≥ ε) = 0.
\]
Let \( m(\delta) = \max\{r | \delta 2^r \leq 1\} \), then, \( \delta 2^{m(\delta)} \leq 1 \) and

\[
m(\delta) \to \infty, \quad \text{as} \quad \delta \to 0.
\]

By Lemma 2.4.4, we have

\[
\sup\{|V_{nk}(t,u,y) - V_{nk}(s,x,y)|; (t,u,y), (s,x,y) \in L_{2p+1}(2^{mn}), |(t,u,y) - (s,x,y)| \leq \delta\}
\]

\[
\leq 4(2p+1) \sum_{r=m(\delta)}^{2p+1} \sum_{\mu=1}^{2} \sup_{(t,u,y) \in L_{2p+1}(2^{mn})} |V_{nk}(t,u,y) - V_{nk}( (t,u,y) + e_{\mu}2^{r} ) |.
\]

Choose \( a \in (0,1) \) such that \( a^{4(p+1)\sqrt{2}} > 1 \). Since

\[
\sum_{r=m(\delta)}^{mn} a^{r-m(\delta)} \frac{\epsilon(1-a)}{4(2p+1)^2} = \frac{\epsilon(1-a^{m(\delta)-1})}{4(2p+1)^2} < \frac{\epsilon}{4(2p+1)^2},
\]

we have

\[
P\left( \sup\{|V_{nk}(t,u,y) - V_{nk}(s,x,y)|; (t,u,y), (s,x,y) \in L_{2p+1}(2^{mn}), |(t,u,y) - (s,x,y)| \leq \delta\} \geq \epsilon\right)
\]

\[
\leq \sum_{r=m(\delta)}^{2p+1} \sum_{\mu=1}^{2} \sum_{j_1=0}^{2^r-1} \sum_{j_\mu=0}^{2^{r-1}} \sum_{j_{2p+1}=0}^{2^r} P\left(|V_{nk}(j) - V_{nk}(j+e_{\mu}2^{r})| > \frac{(1-a)a^{r-m(\delta)}}{4(2p+1)^2} \epsilon\right)
\]
\[
\sum_{r=m(\delta)}^{m} \sum_{\mu=1}^{2^p} \sum_{j_1=0}^{2^r} \ldots \sum_{j_\mu=0}^{2^r-1} \ldots \\
\quad \sum_{j_{2p+1}=0}^{2^r} \left( \frac{4(2p+1)^2}{c(1-a)^{r-m(\delta)}} \right)^{4(p+1)} E[V_{nk}(j) - V_{nk}(j+e_\mu 2^{r-1})]^{4(p+1)}.
\]

(2.4.18)

We notice that if a r.v. \( \xi \) satisfies

\[
E \xi^i = \begin{cases} 
1 & i=0 \\
E\xi & i \geq 1 
\end{cases}
\]

then, for \( l \geq 2 \)

\[
E(\xi - E\xi)^l = \sum_{i=0}^{l} \binom{l}{i} E\xi^i (E\xi)^{l-i} = E\xi \sum_{i=1}^{l} \binom{l}{i} (E\xi)^{l-i} + (E\xi)^l
\]

\[
= E\xi[(1+E\xi)^l - (E\xi)^l + (E\xi)^{l-1}].
\]

Let

\[
\Delta_1 = E[V_{nk}(j) - V_{nk}(j+e_1 2^{r-1})]^{4(p+1)} = E\left[ \sum_{i=1}^{n} c_{ni} (\xi_{ni} - E\xi_{ni}) \right]^{4(p+1)}
\]
where

\[
\xi_{ni} = I(F^{-1}(t) + \xi_{ni}^+ (2u-I)K - \xi_{ni}^- (2y-I)K) < Y_i \leq \frac{F^{-1}(t+2^r) + \xi_{ni}^+ (2u-I)K - \xi_{ni}^- (2y-I)K).}
\]

For \( \xi_{ni} \), we have

\[
|E(\xi_{ni} - E\xi_{ni})|^l = \begin{cases} 
1 & l = 0 \\
0 & l = 1 \\
\leq C E\xi_{ni} & 2 \leq l \leq 4(p+1)
\end{cases}
\]

where \( C \) is a constant. By the choice of \( m_n \), (2.4.15) and (2.4.16), we have

\[
a_n = \max_{1 \leq j \leq p} \max_{1 \leq i \leq n} |\xi_{nij}| \leq 2^{-cr}, \quad \text{for } r \leq m_n.
\]

So,

\[
E\xi_{ni} = F(F^{-1}(t+2^r) + \xi_{ni}^+ (2u-I)K - \xi_{ni}^- (2y-I)K) - F(F^{-1}(t) + \xi_{ni}^+ (2u-I)K - \xi_{ni}^- (2y-I)K)
\]

\[
= t + 2^r + F'(\xi)(\xi_{ni}^+ (2u-I)K - \xi_{ni}^- (2y-I)K) -
\]

\[
- t - F'(\eta)(\xi_{ni}^+ (2u-I)K - \xi_{ni}^- (2y-I)K)
\]
\[ \leq 2^{-r} + M_1 a_n \leq M_2 2^{-c r} \]

where \( M_1, M_2 \) are constants.

Let

\[ N_k = \sum_{i=1}^{n} I(k_i \geq 2), \quad (2.4.19) \]

where \( k_1 + \ldots + k_n = 4(p+1), \ 0 \leq k_i \leq 4(p+1) \). Then, \( N_k \leq 2(p+1) \). So,

\[ \Delta_1 = \sum_{k_1 + \ldots + k_n = 4(p+1)} \binom{4(p+1)}{k_1} \ldots \binom{4(p+1)}{k_n} c_{11}^k \ldots c_{nn}^k E(\xi_n - E_\xi_i)^{k_1} \ldots E(\xi_{ni} - E_\xi_{ni})^{k_n} \]

\[ \leq M_3 \sum_{k_1 + \ldots + k_n = 4(p+1)} \prod_{i=1}^{n} c_{nik}^{k_i} c_{nik}^{k_i} E(\xi_{ni} - E_\xi_{ni})^{k_i} \]

\[ \leq M_4 \ a_n \ 4^{4(p+1) - 2N_k} \ 2^{-c r N_k} \leq M_4 \ 2^{-c r [4(p+1) - N_k]} \]

\[ \leq M_4 \ 2^{-c r 2(p+1)} \]

where \( a \wedge b = \min\{a, b\} \), \( a \vee b = \max\{a, b\} \), and \( M_3, M_4 \) are constants. Therefore,

\[ \sum_{j_1=0}^{2^r-1} \sum_{j_2=0}^{2^r-1} \sum_{j_{2p+1} = 0}^{2^r} \left( \frac{4(2p+1)^2}{\epsilon(1-a)a^{r-m(\delta)}} \right)^{4(p+1)} E[V_{nk}(j) - V_{nk}(j + e_1 2^{-r})]^{4(p+1)} \]

\[ \leq M_4 \sum_{j_1=0}^{2^r-1} \sum_{j_2=0}^{2^r-1} \sum_{j_{2p+1} = 0}^{2^r} \left( \frac{4(2p+1)^2}{\epsilon(1-a)a^{r-m(\delta)}} \right)^{4(p+1)} 2^{-c r 2(p+1)} \]
\[
\leq M_5 \left( \frac{4(2p+1)^2}{\epsilon(1-a) \alpha^{r-m(\delta)}} \right)^{4(p+1)} 2(2p+1)^2 - cr2(p+1)
\]

\[
= M_5 \left( \frac{4(2p+1)^2}{\epsilon(1-a) \alpha^{r-m(\delta)}} \right)^{4(p+1)} 2^{-\frac{r}{2}}
\]

(2.4.20)

where \( M_5 \) is a constant.

Similarly, let

\[
\Delta_\mu = E[V_{nk} (j) - V_{nk} (j + e\mu 2^{-r})]^{4(p+1)}
\]

\[
= E\left[ \sum_{i=1}^{n} c_{nik} (\eta_{ni} - E\eta_{ni}) \right]^{4(p+1)}, \quad 2 \leq \mu \leq p + 1
\]

\[
\Delta_\mu = E[V_{nk} (j) - V_{nk} (j + e\mu 2^{-r})]^{4(p+1)}
\]

\[
= E\left[ \sum_{i=1}^{n} c_{nik} (\gamma_{ni} - E\gamma_{ni}) \right]^{4(p+1)}, \quad p + 2 \leq \mu \leq 2p + 1
\]

where

\[
\eta_{ni} = I\left\{ F^{-1}(t) + (\varepsilon_{ni}^+, -\varepsilon_{ni}^-) \left( (2u - I)^\tau, (2y - I)^\tau \right)^\tau K < Y_i \leq \right.
\]

\[
\leq F^{-1}(t) + (\varepsilon_{ni}^+, -\varepsilon_{ni}^-) \left( (2u - I)^\tau, (2y - I)^\tau \right)^\tau K + e\mu 2^{-r+1} \right); \]

\[
\gamma_{ni} = I\left\{ F^{-1}(t) + (\varepsilon_{ni}^+, -\varepsilon_{ni}^-) \left( (2u - I)^\tau, (2y - I)^\tau \right)^\tau K + e\mu 2^{-r+1} \right) < Y_i \leq \right.
\]
\[
\leq F^{-1}(t) + (\eta_{ni}^+, \eta_{ni}^-) \left( (2y - 1)^\tau, (2y - 1)^\tau \right)^\tau K.
\]

Then,

\[
E_{\eta_{ni}} = F \left\{ F^{-1}(t) + (\eta_{ni}^+, \eta_{ni}^-) \left( (2y - 1)^\tau, (2y - 1)^\tau \right)^\tau K + e_\mu 2^{-r+1} \right\} - F \left\{ F^{-1}(t) + (\eta_{ni}^+, \eta_{ni}^-) \left( (2y - 1)^\tau, (2y - 1)^\tau \right)^\tau K \right\}
\]

\[
= F'(\xi) \ c_{ni\mu}^+ 2^{-r+1} \leq M_6 2^{-cr},
\]

\[
E_{\eta_{ni}} = F \left\{ F^{-1}(t) + (\eta_{ni}^+, \eta_{ni}^-) \left( (2y - 1)^\tau, (2y - 1)^\tau \right)^\tau K \right\} - F \left\{ F^{-1}(t) + (\eta_{ni}^+, \eta_{ni}^-) \left( (2y - 1)^\tau, (2y - 1)^\tau \right)^\tau K + e_\mu 2^{-r+1} \right\}
\]

\[
= F'(\eta) \ c_{ni(\mu-p-1)}^- 2^{-r+1} \leq M_6 2^{-cr},
\]

where \( M_6 \) is a constant. Therefore, for \( 2 \leq \mu \leq 2p+1 \),

\[
\Delta_\mu = E[V_{nk}(j) - V_{nk}(j + e_\mu 2^{-r})]^{4(p+1)} \leq M_7 2^{-cr2(p+1)},
\]

where \( M_7 \) is a constant. In (2.4.18), we have

\[
\frac{4(2p+1)^2}{\kappa(1-a)^{r-m(\delta)}} 4(p+1) E[V_{nk}(j) - V_{nk}(j + e_\mu 2^{-r})]^{4(p+1)}
\]
\[ \leq M_7 \sum_{j_1=0}^{2^r} \ldots \sum_{j_{2p+1}=0}^{2^r} \sum_{j_{2p+2} \cdots} \left( \frac{4(2p+1)^2}{(1-a)^{r-m(\delta)}} \right)^{4(p+1)} 2^{-cr(2p+1)} \]

\[ \leq M_8 \left( \frac{4(2p+1)^2}{(1-a)^{r-m(\delta)}} \right)^{4(p+1)} 2^{-\frac{r}{2}}, \quad (2.4.21) \]

where \(M_8\) is a constant.

From (2.4.18), (2.4.20) and (2.4.21), we have, as \(\delta \to 0\)

\[ P\left( \sup \{|V_{nk}(t, u, y) - V_{nk}(s, x, y)|; \right. \]

\[ (t, u, y), (s, x, y) \in L_{2p+1}(2^{mn}), |(t, u, y) - (s, x, y)| \leq \delta \} \geq \epsilon \}

\[ \leq \sum_{r=m(\delta)}^{\infty} M_9 \left( \frac{4(2p+1)^2}{(1-a)^{r-m(\delta)}} \right)^{4(p+1)} 2^{-\frac{r}{2}} \]

\[ \leq M_9 \left( \frac{4(2p+1)^2}{(1-a)^{r-m(\delta)}} \right)^{4(p+1)} 2^{-\frac{m(\delta)}{2}} \sum_{r=0}^{\infty} \rho^r \to 0 \]

where \(M_9\) is a constant, because \(\rho = (a^{4(p+1)\sqrt{2}})^{-1} < 1\) and \(m(\delta) \to \infty\), as \(\delta \to 0\).

\[ \square \]

**PROPOSITION 2.4.6.** For any fixed \(k=1, 2, \ldots, p\), assume (A1), (A2)

and (B). Let

\[ T_{nk}(t, u) = \tilde{S}_{nk}^* (t, (2u-I)K) - t \sum_{i=1}^{n} c_{nik}, \quad (t, u) \in E_{p+1} \]
and let \( \{P_{nk}; n \geq 1\} \) be the sequence of probability measures corresponding to \( T_{nk}, n \geq 1 \). Then, \( \{P_{nk}\} \) is relatively compact.

Proof. Note that \( T_{nk}(t, y) \in C[0,1]^{p+1} \) and \( S^{*}_{nk}(0, -KI) = 0 \), for \( n \geq 1 \). By (2.4.4), we have

\[
|T_{nk}(0,0)| = |S^{*}_{nk}(0, -KI)| \leq (p+1) \max_{1 \leq i \leq n} c_{nik}, \text{ a.s.}
\]

Hence,

\[
P_{nk}^o = P \pi_0^{-1} \text{ converges in distribution}. \quad (2.4.22)
\]

By virtue of Neuhaus' (1971, discussion on page 1290-1291), it suffices to show that for any \( \epsilon > 0 \),

\[
\lim_{\delta \to 0} \lim_{n \to \infty} P\left( \omega_{T_{nk}}(\delta) \geq \epsilon \right) = 0, \quad (2.4.23)
\]

because (2.4.22) and (2.4.23) imply that \( \{P_{nk}\} \) is relatively compact.

Since for any \( (t, y), (s, y) \in E_{p+1} \),

\[
|T_{nk}(t, y) - T_{nk}(s, y)|
\]

\[
= |[S^{*}_{nk}(t, (2y-I)K) - S^{*}_{nk}(s, (2y-I)K)] - (t-s)\sum_{i=1}^{n} c_{nik}|
\]

\[
\leq \left[ |S^{*}_{nk}(t, (2y-I)K) - S^{*}_{nk}(s, (2y-I)K)| - |S^{*}_{nk}(t, (2y-I)K) - S^{*}_{nk}(s, (2y-I)K)| \right] +
\]
\[ + |W_n(k)(t, u) - W_n(k)(s, y)| + |W^{o}_n(k)(t, u) - W^{o}_n(k)(s, y)| \]

by (2.4.4), we have

\[ \omega_{n,k}(\delta) \leq 2(p+1) \max_{1 \leq i \leq n} c_{n,k} + \omega_{W_n(k)}(\delta) + \omega_{W^{o}_n(k)}(\delta), \quad \text{a.s.} \]

By Corollary 2.4.3, we have (2.4.14) for $W^{o}_{n,k}$ from a similar proof. Therefore, (2.4.23) follows from (2.4.2), (2.4.14) for $W^{o}_{n,k}$ and Proposition 2.4.5. \(\square\)

**Lemma 2.4.7.** Suppose \(\Gamma = \{T_{\lambda}; \lambda \in \Lambda\}\) is a compact set in \(C[0,1]^{p+1}\), let \(\Gamma_1 = \{T_{\lambda}(\cdot, \mu); \lambda \in \Lambda, \mu \in \mathbb{E}_p\}\), then \(\Gamma_1\) is a compact set in \(C[0,1]\).

**Proof.** It suffices to show that any infinite sequence in \(\Gamma_1\) contains a convergent subsequence in \(\Gamma_1\).

Suppose \(\{T_{\lambda_n}(\cdot, \mu_n)\}\) is a sequence of \(\Gamma_1\), then \(\{T_{\lambda_n}(\cdot, \cdot)\}\) is a sequence of \(\Gamma\). Since \(\Gamma\) is compact, \(\{T_{\lambda_n}(\cdot, \cdot)\}\) contains a convergent subsequence also denoted by \(\{T_{\lambda_n}(\cdot, \cdot)\}\) such that

\[ \|T_{\lambda_n} - T_{\lambda_0}\| \to 0, \quad \text{as } n \to \infty \quad (2.4.24) \]

where \(\lambda_0 \in \Lambda\). Since \(\mu_n \in \mathbb{E}_{p+1}\), there exists a convergent subsequence also denoted by \(\mu_n\) such that...
\[ y_n \to y_0 \in E_{p+1}, \quad \text{as } n \to \infty. \quad (2.4.25) \]

For any \( t \in [0,1] \), we have
\[ |T_{\lambda_n}(t,y_n) - T_{\lambda_0}(t,y_0)| \leq |T_{\lambda_n}(t,y_n) - T_{\lambda_0}(t,y_n)| + |T_{\lambda_0}(t,y_n) - T_{\lambda_0}(t,y_0)|. \]

Therefore, by (2.4.24), (2.4.25) and the uniform continuity of \( T_{\lambda_0} \) on \( E_{p+1} \), we have
\[ \|T_{\lambda_n}(\cdot, y_n) - T_{\lambda_0}(\cdot, y_0)\| \to 0, \quad \text{as } n \to \infty \]

where \( T_{\lambda_0}(\cdot, y_0) \in \Gamma_1 \).

PROPOSITION 2.4.8. For any fixed \( k = 1, 2, \ldots, p \), assume (A1), (A2) and (B). If, for any compact set \( \Gamma_0 \) in \( C[0,1] \),
\[ \lim_{t \to 0} \frac{\operatorname{Rem}(tH)}{t} = 0, \quad (2.4.26) \]

uniformly for \( H \in \Gamma_0 \). Then, for any \( K > 0 \), as \( n \to \infty \)
\[ \sup_{|y| \leq K} \left| \sum_{i=1}^{n} c_{nik} \frac{\operatorname{Rem} \left( \frac{S^*_n(\cdot, y)}{\sum_{i=1}^{n} c_{nik}} - U(\cdot) \right)}{P} \right| = 0. \quad (2.4.27) \]
Proof. From Proposition 2.4.6, we know that \( \{ \mathbf{P}_{nk} \} \) is relatively compact in \( \mathbb{P}_{[0,1]}^{p+1} \), where

\[
\mathbf{P}_{nk}(A) = \mathbf{P}( \mathbf{T}_{nk} \in A ).
\]

Since \( \mathbb{P}_{[0,1]}^{p+1} \) is complete and separable, by Prohorov's theorem (Billingsley, 1968, Theorem 6.2), \( \{ \mathbf{P}_{nk} \} \) is tight, i.e., for any \( \varepsilon > 0 \), there exists a compact set \( \Gamma \) in \( \mathbb{P}_{[0,1]}^{p+1} \) such that

\[
\mathbf{P}( \mathbf{T}_{nk} \in \Gamma ) > 1 - \varepsilon, \quad \text{for } n \geq 1. \tag{2.4.28}
\]

Using the same definition about \( \Gamma_1 \) in Lemma 2.4.7 with respect to \( \Gamma \), we know that \( \Gamma_1 \) is a compact set of \( \mathbb{P}_{[0,1]} \). By (2.4.3) and (2.4.26), there exists a positive integer \( N \) such that

\[
\left| \sum_{i=1}^{n} c_{nik} \operatorname{Rem} \left( \frac{H}{\sum_{i=1}^{n} c_{nik}} \right) \right| \leq \varepsilon, \quad \text{for } n \geq N, H \in \Gamma_1. \tag{2.4.29}
\]

If \( \mathbf{T}_{nk} \in \Gamma \), then \( \mathbf{T}_{nk}(\cdot, y) \in \Gamma_1 \), for any \( y \in \mathbb{E}_p \), and, by (2.4.29),

\[
\left| \sum_{i=1}^{n} c_{nik} \operatorname{Rem} \left( \frac{\mathbf{T}_{nk}(\cdot, y)}{\sum_{i=1}^{n} c_{nik}} \right) \right| \leq \varepsilon, \quad \text{for } n \geq N, y \in \mathbb{E}_p. \tag{2.4.30}
\]

Since (2.4.30) implies

\[
\sup_{y \in \mathbb{E}_p} \left| \sum_{i=1}^{n} c_{nik} \operatorname{Rem} \left( \frac{\mathbf{T}_{nk}(\cdot, y)}{\sum_{i=1}^{n} c_{nik}} \right) \right| \leq \varepsilon, \quad \text{for } n \geq N \tag{2.4.31}
\]
and since (2.4.31) can be equivalently written as

\[
\sup_{|u|\leq K} \left| \sum_{i=1}^{n} c_{nik} \operatorname{Rem}\left( \frac{\tilde{S}_{nk}^{*}(\cdot, u)}{\sum_{i=1}^{n} c_{nik}} - U(\cdot) \right) \right| \leq \epsilon, \quad \text{for } n \geq N \tag{2.4.32}
\]

therefore, from (2.4.28) and (2.4.32), we have, for \( n \geq N \),

\[
1 - \epsilon < P(T_{nk} \in \Gamma) \leq P\left\{ \sup_{|u|\leq K} \left| \sum_{i=1}^{n} c_{nik} \operatorname{Rem}\left( \frac{\tilde{S}_{nk}^{*}(\cdot, u)}{\sum_{i=1}^{n} c_{nik}} - U(\cdot) \right) \right| \leq \epsilon \right\}.
\]

\[\square\]

Let \( \Gamma \) be a set in \( D[0,1] \) and \( H \in D[0,1] \), define

\[
\text{dist}(H, \Gamma) = \inf_{G \in \Gamma} \| H - G \|. \tag{2.4.33}
\]

**Lemma 2.4.9.** Let \( Q: D[0,1] \times R \to R \) and suppose that for any compact set \( \Gamma \) in \( D[0,1] \),

\[
\lim_{t \to 0} Q(H, t) = 0 \tag{2.4.34}
\]

uniformly for \( H \in \Gamma \). Let \( \epsilon > 0 \) and let \( \alpha_n, \beta_n \) be sequences of real numbers such that \( \alpha_n \to 0, \beta_n \to 0 \), as \( n \to \infty \). Then, for any compact set \( \Gamma \) in \( D[0,1] \), there exists a positive integer \( N \) such that, if \( \text{dist}(H, \Gamma) \leq \alpha_n \), then
\[ |Q(H, \beta_n)| < \epsilon, \quad \text{for } n \geq N. \]

Proof. Suppose not. Then, for a real number \( \epsilon > 0 \), there exists a compact set \( \Gamma \) in \( D[0,1] \) and sequence \( \{H_k\} \subset D[0,1] \) with \( \text{dist}(H_k, \Gamma) \leq \alpha_n \) such that

\[ |Q(H_k, \beta_n)| \geq \epsilon. \quad (2.4.35) \]

Since \( \text{dist}(H_k, \Gamma) \leq \alpha_n \), we can choose \( H_k^* \in \Gamma \) such that

\[ \|H_k - H_k^*\| \leq \alpha_n. \]

Since \( \{H_k^*\} \subset \Gamma \) and \( \Gamma \) is a compact set, \( \{H_k^*\} \) has an accumulation point \( H^* \in \Gamma \). Therefore, we can choose a subsequence of \( \{H_k^*\} \) also denoted by \( \{H_k^*\} \) such that

\[ H_k^* \to H^*, \quad \text{as } k \to \infty. \]

Since \( \alpha_n \to 0 \), we also have

\[ H_k \to H^*, \quad \text{as } k \to \infty. \]

and the set

\[ \Gamma_1 = \{H_k; k \geq 1\} \cup \{H^*\} \]

is compact. By (2.4.34), we have
\[ \text{Q}(H_k, t) \to 0, \quad \text{as } t \to 0. \]

uniformly for \( H_k \in \Gamma_1 \). This contradicts (2.4.35). \( \square \)

From the discussion by Fernholz (1983), we know that [\( S_{nk}^* (\cdot, u) - U(\cdot) \sum_{i=1}^{n} c_{nik} \)] are not random elements of \( D[0,1] \) with uniform topology. Next, we will use the inner probability measure \( P_* \) corresponding to \( P \) to deal with this problem.

**Lemma 2.4.10.** Under the assumptions of Proposition 2.4.8, for \( \epsilon > 0 \), there exists a compact set \( \Gamma \) in \( D[0,1] \) such that, for all \( n \geq 1 \),

\[ P_* \left\{ \text{dist} \left( S_{nk}^* (\cdot, (2u - 1)K) - U(\cdot) \sum_{i=1}^{n} c_{nik}, \Gamma \right) \leq (p+1) \max_{1 \leq i \leq n} c_{nik}, \forall u \in E_p \right\} > 1 - \epsilon. \]

**Proof.** From the proof of Proposition 2.4.8, there exists a compact set \( \Gamma \) in \( C[0,1]^{p+1} \) such that, for all \( n \geq 1 \),

\[ P(T_{nk} \in \Gamma) > 1 - \epsilon. \]

By (2.4.4), we have, for \( n \geq 1 \),
\[ P\{T_{nk} \in \Gamma, \| \tilde{S}_{nk}(\cdot,z) - S_{nk}(\cdot,z) \| \leq (p+1) \max_{1 \leq i \leq n} c_{nik} \} \geq 1 - \epsilon. \] (2.4.36)

By Lemma 2.4.7, we know that \( \Gamma \) induces a compact set \( \Gamma_1 \) in \( C[0,1] \), which is also a compact set of \( D[0,1] \) because \( C[0,1] \) is a subspace of \( D[0,1] \). We also know that if \( T_{nk} \in \Gamma \), then \( T_{nk}(\cdot,u) \in \Gamma_1 \) for any \( u \in E_p \), i.e.,

\[ [\tilde{S}_{nk}(\cdot,(2u-1)K) - U(\cdot) \sum_{i=1}^{n} c_{nik}] \in \Gamma_1, \quad \text{for any } u \in E_p. \] (2.4.37)

By the definition, we know that \( \| \tilde{S}_{nk}(\cdot,z) - S_{nk}(\cdot,z) \| \leq (p+1) \max_{1 \leq i \leq n} c_{nik} \) implies

\[ \| \tilde{S}_{nk}(\cdot,(2u-1)K) - S_{nk}(\cdot,(2u-1)K) \| \leq (p+1) \max_{1 \leq i \leq n} c_{nik}, \quad \text{for any } u \in E_p. \] (2.4.38)

Since (2.4.37) and (2.4.38) imply

\[ \text{dist}([S_{nk}(\cdot,(2u-1)K) - U(\cdot) \sum_{i=1}^{n} c_{nik}], \Gamma_1) \leq (p+1) \max_{1 \leq i \leq n} c_{nik}, \quad \forall u \in E_p, \]

by (2.4.36), we have, for \( n \geq 1 \),

\[ P_*\{ \text{dist}([S_{nk}(\cdot,(2u-1)K) - U(\cdot) \sum_{i=1}^{n} c_{nik}], \Gamma_1) \leq (p+1) \max_{1 \leq i \leq n} c_{nik}, \forall u \in E_p \} > 1 - \epsilon. \]
Proof of Theorem 2.3.1. Since \( \tau: D[0,1] \rightarrow R \) is Hadamard differentiable at \( U \), by the definition of Hadamard differentiability (2.2.27), (2.4.26) holds.

By Lemma 2.4.10, for \( \epsilon > 0 \), there exists a compact set \( \Gamma \) in \( D[0,1] \) such that, for \( n \geq 1 \),

\[
P_\ast \{ \text{dist}([S^*_{nk} (\cdot, (2u-I)K) - U(\cdot) \sum_{i=1}^{n} c_{nik}], \Gamma) \leq (p+1) \max_{1 \leq i \leq n} c_{nik}, \forall u \in E_p \} > 1 - \frac{\epsilon}{2}.
\]

Therefore, we can find measurable sets \( B_n \) for all \( n \) such that

\[
B_n \subset \{ \text{dist}([S^*_{nk} (\cdot, (2u-I)K) - U(\cdot) \sum_{i=1}^{n} c_{nik}], \Gamma) \leq (p+1) \max_{1 \leq i \leq n} c_{nik}, \forall u \in E_p \}
\]

and

\[
P(B_n) > 1 - \epsilon. \quad (2.4.39)
\]

Using Lemma 2.4.9 by considering \( Q(H,t) = \text{Rem}(tH)/t \), for the compact set \( \Gamma \), there exists a positive integer \( N \) such that, if \( \text{dist}(H, \Gamma) \leq (p+1) \max_{1 \leq i \leq n} c_{nik} \) for \( n \geq N \), then

\[
\left| \sum_{i=1}^{n} c_{nik} \text{Rem}(\frac{H}{\sum_{i=1}^{n} c_{nik}}) \right| < \epsilon.
\]

Therefore, taking \( H = [S^*_{nk} (\cdot, (2u-I)K) - U(\cdot) \sum_{i=1}^{n} c_{nik}] \) for \( n \geq N \) and any \( u \in E_p \),

\[
\{ \text{dist}([S^*_{nk} (\cdot, (2u-I)K) - U(\cdot) \sum_{i=1}^{n} c_{nik}], \Gamma) \leq (p+1) \max_{1 \leq i \leq n} c_{nik}, \forall u \in E_p \}
\]

implies
\[ \left| \sum_{i=1}^{n} c_{n_{ik}} \text{Rem} \left( \frac{S_{nk}^* (\cdot, y) - U(\cdot)}{\sum_{i=1}^{n} c_{n_{ik}}} \right) \right| < \epsilon, \quad \text{for } y \in E_p. \] (2.4.40)

Since (2.4.40) implies

\[ \sup_{|y| \leq K} \left| \sum_{i=1}^{n} c_{n_{ik}} \text{Rem} \left( \frac{S_{nk}^* (\cdot, y) - U(\cdot)}{\sum_{i=1}^{n} c_{n_{ik}}} \right) \right| \leq \epsilon, \] (2.4.41)

by (2.4.39), we have, for \( n \geq N, \)

\[ 1 - \epsilon < P(B_n) \leq P \{ \sup_{|y| \leq K} \left| \sum_{i=1}^{n} c_{n_{ik}} \text{Rem} \left( \frac{S_{nk}^* (\cdot, y) - U(\cdot)}{\sum_{i=1}^{n} c_{n_{ik}}} \right) \right| \leq \epsilon \}. \]

\[ \square \]

**Remark (1).** In the proof of Theorem 2.3.1, we used the fact that, for any \( n \geq 1 \) and \( 1 \leq k \leq p, \)

\[ \text{Rem} \left( \frac{S_{nk}^* (\cdot, y) - U(\cdot)}{\sum_{i=1}^{n} c_{n_{ik}}} \right), \quad y \in \mathbb{R}^p \]

is measurable, even though \( S_{nk}^* (\cdot, y) \) is not a random element of \( D[0,1], \) and therefore we used probability measure \( P \) for events concerning this function. This is because that

\[ r \left( \frac{S_{nk}^* (\cdot, y)}{\sum_{i=1}^{n} c_{n_{ik}}} \right) = T \left( \frac{S_{nk}^* (F(\cdot), y)}{\sum_{i=1}^{n} c_{n_{ik}}} \right) \]
is measurable, and that, by Lemma 4.4.1 of Fernholz (1983), we know that

\[
\tau_U(S_{nk}^*(\cdot, u) - U(\cdot) \sum_{i=1}^n c_{nik}) = \sum_{i=1}^n c_{nik} \tau_U((\delta_{(y_i - c_{ni}^T u)} - F) \circ F^{-1})
\]

\[
= \sum_{i=1}^n c_{nik} IC(Y_i - c_{ni}^T u; F, T), \tag{2.4.42}
\]

is also measurable, where (2.4.42), without any assumptions on \( \{c_{nik}\} \) and the underlying d.f. \( F \), is true as long as \( \tau \) is Hadamard differentiable at \( U \).

**Remark (2).** If in Theorem 2.3.1 and Proposition 2.4.8,

\[
\sup_{|u| \leq K} \left| \sum_{i=1}^n c_{nik} \text{Rem} \left( \frac{S_{nk}^*(\cdot, u)}{\sum_{i=1}^n c_{nik}} - U(\cdot) \right) \right|
\]

or

\[
\sup_{|u| \leq K} \left| \sum_{i=1}^n c_{nik} \text{Rem} \left( \frac{S_{nk}^*(\cdot, u)}{\sum_{i=1}^n c_{nik}} - U(\cdot) \right) \right|
\]

is not measurable, we replace \( |u| \leq K \) by \( u \in Q_K^P \), where

\[
Q_K^P = \{ \text{all rational points in } [-K, K]^P \}.
\]

Then, they both are measurable. Our main results in Section 2.3 will be slightly different, but still good enough for the study of (2.1.15).
CHAPTER III

ON M-ESTIMATION IN LINEAR MODELS

3.1 Introduction

The linear regression model is one of the most widely used tools in statistical analysis. In this chapter, we consider the simple linear regression model as the following:

\[ X_i = \beta^T c_i + e_i, \quad i \geq 1 \]  \hspace{1cm} (3.1.1)

where the \( c_i \) are known \( p \)-vectors of regression constants, \( \beta=(\beta_1, \ldots, \beta_p)^T \) is the vector of unknown (regression) parameters, \( p \geq 1 \), and is to be estimated from \( n \) observations \( X_1, \ldots, X_n \), and \( e_i \) are independent and identically distributed random variables (i.i.d.r.v.) with distribution function (d.f.) \( F \). We should notice that (3.1.1) reduces to the classical location model if all \( c_i \), with \( p=1 \), are equal to 1.

Classically, the problem is solved by minimizing the sum of squares (with respect to \( \theta \)):

\[ \sum_{i=1}^{n} (X_i - c_i^T \theta)^2 = \min! \]  \hspace{1cm} (3.1.2)
or, equivalently, by solving the system of $p$ equations (with respect to $\theta$) obtained by differentiating (3.1.2)

$$\sum_{i=1}^{n} \epsilon_i (X_i - \gamma_i^T \theta) = 0. \quad (3.1.3)$$

The computation of the Least Square Estimators (LSE) defined by the estimating equations (3.1.3), which are linear equations, can be easily done. Furthermore, the statistical properties of LSE, such as the consistency and the asymptotic normality, have been studied in detail (Huber, 1981). But, as we know, in spite of its mathematical beauty and computational simplicity, the least square estimators suffer a dramatic lack of robustness. Indeed, one single outlier can have an arbitrarily large effect on the estimate. Hence, other ‘robust’ estimators of $\theta$ have been proposed. In particular, Huber (1973) robustized the classical equations (3.1.2) and (3.1.3) in a straightforward way: instead of minimizing a sum of squares, we minimize a sum of a less rapidly increasing function $\rho$ of the residuals, i.e.,

$$\sum_{i=1}^{n} \rho(X_i - \gamma_i^T \theta) = \min! \quad (3.1.4)$$

or, if $\rho$ is convex with derivative $\psi$, (3.1.4) is equivalent to

$$\sum_{i=1}^{n} \epsilon_i \psi(X_i - \gamma_i^T \theta) = 0. \quad (3.1.5)$$

The resulting estimators of (3.1.5), $\hat{\theta}_n$, are called $M$-estimators of Regression and $\psi$ is called Score Function of the M-estimators.
To study the asymptotic properties of the M-estimators of regression, various procedures have been considered by various people.

Huber (1973) proved the asymptotic normality of the M-estimators of regression, assuming that the score function $\psi$ is continuous, nondecreasing and bounded with a bounded second derivative. Koul (1977) and Jurečková (1977, indirectly) also proved the asymptotic normality under weak conditions on $\psi$, but they both required that the underlying d.f. $F$ has finite Fisher information, i.e.,

$$0 < I(f) = \int (f'/f)^2 \, dF < \infty$$

where $F'=f$ is the density function of $F$. As a corollary of Theorem 4.1 of Bickel (1975), Yohai and Maronna (1979) obtained the asymptotic normality under certain assumptions on $\psi$ and $F$, one of which is given as below:

$$\frac{\psi(u+z)}{z} \psi(u) \geq d, \quad \text{if } |u| \leq c \quad |z| \leq b$$  \hspace{1cm} (3.1.6)

where $b$, $c$, and $d$ are positive real numbers satisfying $q=F(c)-F(-c)>0$.

Relles (1968) proved the consistency of this M-estimators of regression when $\psi$ belongs to the Huber family

$$\psi(x, k) = \min\{|x|, k\} \text{sgn}(x).$$

Yohai and Maronna (1979) proved the weak consistency for nondecreasing $\psi$ when (3.1.6) is assumed.

For non-convex $\rho$, Jurečková (1989) derived the asymptotic normality and
the weak consistency of the M-estimators in linear models, assuming certain existence of the second derivative of $F$.

The interest of our current research is to study the asymptotic properties of the M-estimators of regression under the weak conditions on the score function $\psi$ and the underlying d.f. $F$. Since M-estimators are implicitly defined by the estimating equations, it is difficult to study their statistical properties directly. The principle difficulty is the non-linearity of the estimators. Therefore, one of the approaches presented in this chapter is to use Hadamard differentiability through Jurečková- (1971, 1977) uniform asymptotic linearity of the M-estimators of regression. Other approaches are also considered in order to study the strong consistency, to which the Hadamard differentiability approach may not apply. The results show that Hadamard differentiability provides a good method for the study of the asymptotic properties through the linear approximation of the estimator.

Some notations along with basic assumptions and some preliminary lemmas are presented in Section 3.2. The asymptotic normality of the M-estimators of regression is considered in Section 3.3, where none of the assumptions, such as bounded second derivative on $\psi$, or finite Fisher information on $F$, or (3.1.6), or any existence of the second derivative of $F$, is required. The weak and strong consistency of the M-estimators in linear models are investigated in Section 3.4, where neither (3.1.6) nor any existence of the second derivative of $F$ is assumed.
3.2 Preliminaries

For the simple linear model (3.1.1), we consider the following normalized version of the estimating equations in (3.1.5):

\[
M_n(u) = \sum_{i=1}^{n} c_{ni} \psi(Y_i - c_{ni}^T u)
\]  

(3.2.1)

where \( Y_i = X_i - \beta^T c_i \) are i.i.d.r.v.'s with d.f. F, and for every \( n \geq p \),

\[
C_n = \sum_{i=1}^{n} c_i c_i^T = (r_{nij})_{1 \leq i, j \leq p}
\]  

(3.2.2)

\[
C_n^0 = \text{Diag}(\sqrt{r_{n11}}, \ldots, \sqrt{r_{nnp}})
\]  

(3.2.3)

\[
\varepsilon_{ni} = (C_n^0)^{-1} c_i = (c_{ni1}, \ldots, c_{nip})^T, \quad 1 \leq i \leq n
\]  

(3.2.4)

\[
Q_n = \sum_{i=1}^{n} \varepsilon_{ni} c_{ni}^T = (C_n^0)^{-1} C_n (C_n^0)^{-1}
\]  

(3.2.5)

\[
y = C_n^0 (\theta - \beta) \in \mathbb{R}^p.
\]  

(3.2.6)

Then, we see that (3.1.5) is equivalent to

\[
M_n(u) = 0 \quad \text{(with respect to } u),
\]  

(3.2.7)

i.e., \( \hat{u}_n = C_n^0 (\hat{\beta}_n - \beta) \) is a solution of (3.2.7), and that for any \( 1 \leq j \leq p \),

\[
\| \varepsilon_{n,j} \|^2 = 1
\]  

(3.2.8)

where \( \| \cdot \| \) stands for the Euclidean norm and

\[
\varepsilon_{n,j} = (c_{n1j}, \ldots, c_{nnj})^T.
\]  

(3.2.9)
Since $M_n(y)$ involves the following empirical function

$$S_n^*(t,y) = \sum_{i=1}^{n} c_{ni} I(Y_i \leq F^{-1}(t) + c_{ni}^T y), \quad t \in [0,1], \; y \in \mathbb{R}^p$$ (3.2.10)

particularly, for each $y \in \mathbb{R}^P$ and $1 \leq k \leq p$, $M_{nk}(y)$ (the k-th component of $M_n(y)$) is a linear functional of

$$S_{nk}^*(t,y) = \sum_{i=1}^{n} c_{nik} I(Y_i \leq F^{-1}(t) + c_{ni}^T y)$$ (3.2.11)

(the k-th component of $S_n^*(t,y)$), viz.,

$$M_{nk}(t,y) = \int \psi(F^{-1}(t)) \; dS_{nk}^*(t,y),$$ (3.2.12)

we consider the expected value of $S_{nk}^*(t,y)$:

$$S_{nk}(t,y) = \sum_{i=1}^{n} c_{nik} F(F^{-1}(t) + c_{ni}^T y).$$ (3.2.13)

We also write for every $1 \leq i \leq n$, $1 \leq k \leq p$, $t \in [0,1]$ and $y \in \mathbb{R}^P$,

$$c_{nik} = c_{nik}^+ - c_{nik}^-; \quad c_{nik}^+ = \max\{0, c_{nik}\}, \quad c_{nik}^- = \min\{0, c_{nik}\};$$ (3.2.14)

$$\varepsilon_{ni} = \varepsilon_{ni}^+ - \varepsilon_{ni}^-; \quad \varepsilon_{ni}^+ = (c_{ni1}^+, \ldots, c_{nip}^+)^T, \quad \varepsilon_{ni}^- = (c_{ni1}^-, \ldots, c_{nip}^-)^T;$$ (3.2.15)

$$\varepsilon_{n,k} = \varepsilon_{n,k}^+ - \varepsilon_{n,k}^-; \quad \varepsilon_{n,k}^+ = (c_{n1k}^+, \ldots, c_{nnk}^+)^T, \quad \varepsilon_{n,k}^- = (c_{n1k}^-, \ldots, c_{nnk}^-)^T;$$ (3.2.16)
\[(d^+_{nk})^2 = \|\varepsilon^+_{n,k}\|^2, \quad (d^-_{nk})^2 = \|\varepsilon^-_{n,k}\|^2; \quad (3.2.17)\]

\[
\varepsilon^+_{nik} = \begin{cases} 
\frac{c^+_{nik}}{d^+_{nk}} & \text{if } d^+_{nk} > 0 \\
0 & \text{if } d^+_{nk} = 0,
\end{cases}
\quad \varepsilon^-_{nik} = \begin{cases} 
\frac{c^-_{nik}}{d^-_{nk}} & \text{if } d^-_{nk} > 0 \\
0 & \text{if } d^-_{nk} = 0;
\end{cases} \quad (3.2.18)
\]

\[
\varepsilon^+_{ni} = (\varepsilon^+_{n1}, \ldots, \varepsilon^+_{nip})^\tau, \quad \varepsilon^-_{ni} = (\varepsilon^-_{n1}, \ldots, \varepsilon^-_{nip})^\tau; \quad (3.2.19)
\]

\[
S^+_{nk}(t,u) = \sum_{i=1}^{n} c^+_{nik} I(Y_i \leq F^{-1}(t) + \varepsilon^\tau_{ni}u),
\]

\[
S^+_{nk}(t,u) = E\{S^+_{nk}(t,u)\}; \quad (3.2.20)
\]

\[
S^-_{nk}(t,u) = \sum_{i=1}^{n} c^-_{nik} I(Y_i \leq F^{-1}(t) + \varepsilon^\tau_{ni}u),
\]

\[
S^-_{nk}(t,u) = E\{S^-_{nk}(t,u)\}; \quad (3.2.21)
\]

so that

\[
S^*_{nk}(t,u) = S^+_{nk}(t,u) - S^-_{nk}(t,u), \quad (3.2.22)
\]

and

\[
\|\varepsilon_{n,k}\|^2 = (d^+_{nk})^2 + (d^-_{nk})^2 = 1. \quad (3.2.23)
\]

Let
\[ \mathbf{I} = (1, \ldots, 1)^T, \quad (3.2.24) \]

and

\[ \mathbf{E}_q = [0, 1]^q \quad (3.2.25) \]

where \( q \) is an arbitrary positive integer, then, for any \( 0 < K \leq \infty, (t, \mathbf{u}) \in \mathbf{E}_{p+1}, (t, \mathbf{u}, \mathbf{y}) \in \mathbf{E}_{2p+1}, 1 \leq k \leq p \), we write

\[ \begin{align*}
\mathbf{W}_{nk}(t, \mathbf{u}) &= S_{nk}^*(t, (2\mathbf{u} - \mathbf{I})K) - S_{nk}(t, (2\mathbf{u} - \mathbf{I})K) \\
\mathbf{W}_{nk}^0(t, \mathbf{u}) &= S_{nk}(t, (2\mathbf{u} - \mathbf{I})K) - S_{nk}(t, 0) \\
\mathbf{G}_{nk}(t, \mathbf{u}) &= \sum_{i=1}^{n} c_{ni} \xi_{ni}^T(2\mathbf{u} - \mathbf{I})F'(F^{-1}(t))K \\
\mathbf{A}_{nk}^*(t, \mathbf{u}, \mathbf{y}) &= \sum_{i=1}^{n} c_{ni} I\left( Y_i \leq F^{-1}(t) + \xi_{ni}^T(2\mathbf{u} - \mathbf{I})K - \xi_{ni}^T(2\mathbf{y} - \mathbf{I})K \right) \\
\mathbf{A}_{nk}(t, \mathbf{u}, \mathbf{y}) &= \mathbb{E}\{\mathbf{A}_{nk}^*(t, \mathbf{u}, \mathbf{y})\} \\
\mathbf{V}_{nk}(t, \mathbf{u}, \mathbf{y}) &= \mathbf{A}_{nk}^*(t, \mathbf{u}, \mathbf{y}) - \mathbf{A}_{nk}(t, \mathbf{u}, \mathbf{y}) \\
\mathbf{V}_{nk}^0(t, \mathbf{u}, \mathbf{y}) &= \mathbf{A}_{nk}(t, \mathbf{u}, \mathbf{y}) - \mathbf{A}_{nk}(t, \frac{1}{2}, \frac{1}{2}) \\
\mathbf{H}_{nk}(t, \mathbf{u}, \mathbf{y}) &= \sum_{i=1}^{n} c_{ni} [\xi_{ni}^T(2\mathbf{u} - \mathbf{I}) - \xi_{ni}^T(2\mathbf{y} - \mathbf{I})] F'(F^{-1}(t))K.
\end{align*} \]
It is easy to see that

\[ W_{nk}(t,u) = V_{nk}(t,u,u) = [A^*_{nk}(t,u,y) - A_{nk}(t,y,y)] \] (3.2.34)

where

\[
\begin{align*}
A^*_{nk}(t,u,y) &= S^*_{nk}(t,(2u-I)K) \\
A_{nk}(t,u,y) &= E\{A^*_{nk}(t,u,y)\} = S_{nk}(t,(2u-I)K)
\end{align*}
\] (3.2.35)

Also, let \( f \) be a function defined on \( E_q \), we denote

\[
\omega_f(\delta) = \sup_{|x-y| \leq \delta} |f(x) - f(y)|
\] (3.2.36)

where \(|x| = |(x_1, \ldots, x_q)| = \max_{1 \leq i \leq q} |x_i|\).

In this chapter, we will always consider the \( D[0,1] \) space (of right continuous real valued functions with left hand limits) is endowed with the uniform topology. Then, the space \( C[0,1] \) of real valued continuous functions, endowed with the uniform topology, is a subspace of \( D[0,1] \). It is easy to see that, for every \( y \in R^p \) and \( 1 \leq k \leq p \), \( S^*_{nk}(\cdot, y) \) is an element of \( D[0,1] \). A convention which we will follow through the whole chapter is: a function \( f: R \to R \) is nondecreasing if \( f(x) \leq f(y) \) for \( x \leq y \), and is increasing if \( f(x) < f(y) \) for \( x < y \). Analogously, for nonincreasing and decreasing.
Some assumptions, which may be required for our results in this chapter, are given as below:

\[(A1) \ c_{nik} \geq 0, \quad i=1, 2, \ldots, n\]

\[(A2) \ \lim_{n \to \infty} n \max_{1 \leq i \leq n} \|c_{n1i}\|^2 < \infty\]

\[(A3) \ \lim_{n \to \infty} n \max_{1 \leq i \leq n} \|c_{n1i}^+\|^2 < \infty, \quad \lim_{n \to \infty} n \max_{1 \leq i \leq n} \|c_{n1i}^-\|^2 < \infty\]

\[(A4) \text{There exists a positive definite } p \times p \text{ matrix } Q \text{ such that } \lim_{n \to \infty} Q_n = Q\]

\[(B) \ F \text{ is absolutely continuous with a derivative } F' \text{ which is positive and continuous with limits at } \pm \infty\]

\[(C1) \ \psi \text{ is a nondecreasing function with a range including positive and negative real numbers}\]

\[(C2) \ f \ \psi \ dF = 0\]

\[(C3) \ 0 < \gamma = f \ \psi' \ dF < \infty\]

\[(C4) \ 0 < \sigma^2 = f \ \psi^2 \ dF < \infty\]

\[(C5) \ \psi \text{ is right continuous and bounded}.\]
We notice that, for any fixed $k=1, \ldots, p$, (A1) and (A2) imply that

\[ \sum_{i=1}^{n} c_{nik} \to \infty, \quad \text{as } n \to \infty, \quad (3.2.37) \]

and (A2) implies two facts: there exists $C > 0$ such that for $n \geq 1$ and any $1 \leq j \leq p$,

\[ \left( \sum_{i=1}^{n} |c_{nij}| \right) a_n \leq C, \quad (3.2.38) \]

where

\[ a_n = \max_{1 \leq j \leq p} \max_{1 \leq i \leq n} |c_{nij}|, \quad (3.2.39) \]

and

\[ \max_{1 \leq i \leq n} c_{nij}^2 \to 0, \quad \text{as } n \to \infty. \quad (3.2.40) \]

We also notice that (B) implies that $F'$ is bounded and uniformly continuous.

For $\bar{x}=(x_1, \ldots, x_q), \bar{y}=(y_1, \ldots, y_q) \in E_q$, we denote

\[ \bar{x} \leq \bar{y} \quad \text{iff} \quad x_i \leq y_i, \quad \text{for } i=1, 2, \ldots, q \quad (3.2.41) \]

**LEMMA 3.2.1.** For a fixed $k=1, \ldots, p$, assume (A1), (A2) and (B). Then, for any $K > 0$, as $n \to \infty$

\[ \sup_{(t,u) \in E_{p+1}} |W_{nk}(t,u) - W_{nk}(t_{i+1})| \to 0. \quad (3.2.42) \]
Proof. Let \( L_q(r) = \{ l_1, \ldots, l_q \mid l_i = 0, 1, \ldots, r; i = 1, 2, \ldots, q \} \), where \( r \) is an arbitrary positive integer. For any \((t, y) \in E_{p+1}\), there exist \((t', y')\), \((t'', y'')\) \(\in L_{p+1}(2^m)\), where \( m \) is an arbitrary positive integer, such that

\[
(t', y') \preceq (t, y) \preceq (t'', y'')
\]

and

\[
|(t', y') - (t, y)| \leq 2^{-m}, \quad |(t'', y'') - (t, y)| \leq 2^{-m}.
\]

We write

\[
W_{n_k}(t, y) - W_{n_k}(t, \frac{1}{2} I) = [W_{n_k}(t, y) - W_{n_k}(t', y')] +
\]

\[
+ [W_{n_k}(t', y') - W_{n_k}(t', \frac{1}{2} I)] + [W_{n_k}(t', \frac{1}{2} I) - W_{n_k}(t, \frac{1}{2} I)].
\]

Consider \([W_{n_k}(t, y) - W_{n_k}(t', y')]\). Since \( A_{n_k}^*(t, u, y) \) is monotonic in each component of \((t, u, y) \in E_{2p+1}\), where \( u, y \in E_p \), from the proof of Proposition 2.4.5, we have

\[
|W_{n_k}(t, y) - W_{n_k}(t', y')| \leq 2^{-m} \sum_{i=1}^{n} c_{n_k} + \sup \{|V_{n_k}(t, u, y) - V_{n_k}(s, x, y)|; (t, u, y), (s, x, y) \in L_{p+1}(2^m), |(t, u, y) - (s, x, y)| \leq 2^{-m+1}\} +
\]

\[
+ \sup \{|V_{n_k}^o(t, u, y) - V_{n_k}^o(s, x, y)|; |(t, u, y) - (s, x, y)| \leq 2^{-m}\}.
\]
Hence, we have

\[
\sup_{(t,y) \in E_{p+1}} |W_{nk}(t,y) - W_{nk}(t,\frac{1}{2}I)| \leq 2^{-m+1} \sum_{i=1}^{n} c_{nik} + \\
+ 2 \sup \{|V_{nk}(t,u,y) - V_{nk}(s,x,y)|; \\
(t,u,y),(s,x,y) \in L_{2p+1}(2^m), |(t,u,y) - (s,x,y)| \leq 2^{-m+1}\} + \\
+ 2 \sup \{|V_{nk}^{o}(t,u,y) - V_{nk}^{o}(s,x,y)|; |(t,u,y) - (s,x,y)| \leq 2^{-m}\} + \\
+ \sup_{(t,u) \in L_{p+1}(2^m)} |W_{nk}(t,u) - W_{nk}(t,\frac{1}{2}I)| \\
\leq 2^{m} v_{nk}^{o} + 2^{-m+1} \sum_{i=1}^{n} c_{nik} + \\
+ \sup_{(t,u) \in L_{p+1}(2^m)} |W_{nk}(t,u) - W_{nk}(t,\frac{1}{2}I)| + \\
+ 2 \sup \{|V_{nk}(t,u,y) - V_{nk}(s,x,y)|; \\
(t,u,y),(s,x,y) \in L_{2p+1}(2^m), |(t,u,y) - (s,x,y)| \leq 2^{-m+1}\}. \tag{3.2.43}
\]

Let

\[
m_n = \max\{m; \frac{1}{a_n} \geq 2^{cm}\}, \text{ where } c = (4p+3)/(4p+4), \tag{3.2.44}
\]

then, by (2.4.14), (2.4.17) and the proof of Proposition 2.4.5, i.e., as \(n \to \infty\)
\[
\sup\{|V_{nk}(t,y,y) - V_{nk}(s,x,y)|; \quad (t,y,y), (s,x,y) \in E_{p+1}(2^{2m})\}, \quad |(t,y,y) - (s,x,y)| \leq 2^{-m+1} \xrightarrow{P} 0 \]

it suffices to show that, as \( n \to \infty \)

\[
\sup_{(t,y) \in E_{p+1}(2^{2m})} |W_{nk}(t,y) - W_{nk}(t, x)| \xrightarrow{P} 0. \tag{3.2.45}
\]

For any \((t,y) \in E_{p+1}\), we write

\[
W_{nk}(t,y) - W_{nk}(t, x) = \sum_{i=1}^{n} c_{nik}(\xi_{ni} - E\xi_{ni}),
\]

where

\[
\xi_{ni} = I(Y_i \leq F^{-1}(y)) + \xi_{ni}^T(2y - 1)K - I(Y_i \leq F^{-1}(y)).
\]

We notice that if a r.v. \( \xi \) satisfies

\[
E(|\xi|^j) = \begin{cases} 
1 & j = 0 \\
E|\xi| & j \geq 1 
\end{cases}
\]

then, for \( l \geq 2 \)

\[
|E(\xi - E\xi)^l| = \left| \sum_{i=0}^{l} \binom{l}{i} E\xi^i (E\xi)^{l-i} \right|
\]
\[
\leq E|\xi| \sum_{i=1}^{l} \binom{l}{i} (E|\xi|)^{l-i} + (E|\xi|)^{l}
\]

\[
= E|\xi| \{[1+E|\xi|]^{l}-[E|\xi|]^{l}+[E|\xi|]^{l-1}\}.
\]

So, for \(\xi_{ni}\), we have

\[
|E(\xi_{ni}-E\xi_{ni})|^{l} = \begin{cases} 
1 & l=0 \\
0 & l=1 \\
\leq M E|\xi_{ni}| & 2 \leq l \leq 4(p+1)
\end{cases}
\]

where \(M\) is a constant. Since,

\[
E|\xi_{ni}| = |F(F^{-1}(t)+\xi_{ni}^{T}(2y-I)K)-F(F^{-1}(t))|
\]

\[
= |F'(\eta) \xi_{ni}^{T}(2y-I)K| \leq M_{1}K_{\eta}
\]

where \(\eta\) is between \([F^{-1}(t)+\xi_{ni}^{T}(2y-I)K]\) and \(F^{-1}(t)\), and \(M_{1}\) is a constant, for any \(\rho>0\), from the proof of Proposition 2.4.5, we have

\[
P\{ \sup_{(t,u)\in L_{p+1}(2^{m_{n}})} |W_{nk}(t,u)-W_{nk}(t,\frac{1}{2}I)|>\rho \}
\]

\[
\leq (2^{m_{n}+1})^{p+1} P\{ |W_{nk}(t,u)-W_{nk}(t,\frac{1}{2}I)|>\rho \}
\]

\[
\leq \frac{(2^{m_{n}+1})^{p+1}}{\rho^{4(p+1)}} E[W_{nk}(t,u)-W_{nk}(t,\frac{1}{2}I)]^{4(p+1)}
\]
\[
\frac{(2^{m_n+1})^{p+1}}{\rho^{4(p+1)}} \mathbb{E}[\sum_{i=1}^{n} c_{ni}(\xi_{ni} - E\xi_{ni})]^{4(p+1)} \leq \frac{2^{m_n+1}}{\rho^{4(p+1)}} M_2 a_n^{2(p+1)} \leq \frac{M_3 2^{m_n(p+1)}}{\rho^{4(p+1)}} a_n^{2(p+1)},
\]

where \(M_2\) and \(M_3\) are constants. By the choice of \(m_n\) given by (3.2.44), (3.2.45) follows from the following:

\[
2^{m_n(p+1)} a_n^{2(p+1)} \leq 2^{m_n(p+1)} 2^{-cm_n2(p+1)}
\]

\[
= 2^{\frac{m_n(2p+1)}{2}} \to 0, \quad \text{as } n \to \infty
\]

because \(m_n \to \infty\), as \(n \to \infty\).

\[\square\]

**COROLLARY 3.2.2.** For a fixed \(k=1, \ldots, p\), assume (A1), (A2) and (B). Then, for any \(K>0\), as \(n \to \infty\)

\[
\sup \left\{ \left| \left[ S_{nk}^*(t,u) - S_{nk}^*(t,\theta) \right] - \sum_{i=1}^{n} c_{ni} c_{ni}^T u F'(F^{-1}(t)) \right| ; \ t \in [0,1], |u| \leq K \right\} \mathbb{P} 0.
\]

(3.2.46)

Proof. Lemma 3.2.1 implies that, for any \(K>0\), as \(n \to \infty\).
\[
\sup \{|[S_{nk}^*(t, y) - S_{nk}^*(t, 0)] - [S_{nk}(t, y) - S_{nk}(t, 0)]|; \ t \in [0, 1], \ |y| \leq K\} \overset{P}{\to} 0.
\] (3.2.47)

Therefore, (3.2.46) follows from Corollary 2.4.3 and (3.2.47).

\[\square\]

**COROLLARY 3.2.3.** Assume (A3) and (B). Then, for any \(K > 0\) and \(1 \leq k \leq p\), as \(n \to \infty\),

\[
\sup \{|[S_{nk}^{*+}(t, y) - S_{nk}^{*+}(t, 0)] - [S_{nk}^+(t, y) - S_{nk}^+(t, 0)]|; \ t \in [0, 1], \ |y| \leq K\} \overset{P}{\to} 0,
\] (3.2.48)

and

\[
\sup \{|[S_{nk}^{*-}(t, y) - S_{nk}^{*-}(t, 0)] - [S_{nk}^-(t, y) - S_{nk}^-(t, 0)]|; \ t \in [0, 1], \ |y| \leq K\} \overset{P}{\to} 0.
\] (3.2.49)

Therefore,

\[
\sup \{|[S_{nk}^{*+}(t, y) - S_{nk}^{*+}(t, 0)] - \sum_{i=1}^{p} c_{niki}^+ s_{ni}^T y F'(F^{-1}(t))|; \ t \in [0, 1], \ |y| \leq K\} \overset{P}{\to} 0,
\] (3.2.50)

and

\[
\sup \{|[S_{nk}^{*-}(t, y) - S_{nk}^{*-}(t, 0)] - \sum_{i=1}^{p} c_{niki}^- s_{ni}^T y F'(F^{-1}(t))|; \ t \in [0, 1], \ |y| \leq K\} \overset{P}{\to} 0.
\] (3.2.51)
Proof. Consider any $1 \leq k \leq p$. If $d^+_{nk} = 0$ or $d^-_{nk} = 0$, it is the case of Corollary 3.2.2. If $d^+_{nk} > 0$ and $d^-_{nk} > 0$, by Lemma 3.2.1, we have, as $n \to \infty$

\[
\frac{1}{d^+_{nk}} \sup_{t \in [0,1], |y| \leq K} \left\{ |S^+_{nk}(t,y) - S^+_{nk}(t,0)| - |S^+_{nk}(t,y) - S^+_{nk}(t,0)| \right\} \overset{P}{\to} 0,
\]

and

\[
\frac{1}{d^-_{nk}} \sup_{t \in [0,1], |y| \leq K} \left\{ |S^-_{nk}(t,y) - S^-_{nk}(t,0)| - |S^-_{nk}(t,y) - S^-_{nk}(t,0)| \right\} \overset{P}{\to} 0.
\]

(3.2.52)

Hence, (3.2.48) and (3.2.49) follow from the fact: $0 < d^+_{nk} < d^-_{nk} < 1$. Further, from the similar proof of Corollary 2.4.3, we have, as $n \to \infty$

\[
\sup_{t \in [0,1], |y| \leq K} \left\{ |S^+_{nk}(t,y) - S^+_{nk}(t,0)| - \sum_{i=1}^{n} c^+_{nik} \xi^+_{ni} y F'(F^{-1}(t)) \right\} \overset{P}{\to} 0,
\]

and

\[
\sup_{t \in [0,1], |y| \leq K} \left\{ |S^-_{nk}(t,y) - S^-_{nk}(t,0)| - \sum_{i=1}^{n} c^-_{nik} \xi^-_{ni} y F'(F^{-1}(t)) \right\} \overset{P}{\to} 0.
\]

(3.2.54)

(3.2.55)

Therefore, (3.2.50) and (3.2.51) follow from (3.2.48), (3.2.54) and (3.2.49), (3.2.55), respectively.

\[\Box\]
LEMMA 3.2.4. Assume that \( \max_{1 \leq i \leq n} c_{nik}^{2} \to 0 \), as \( n \to \infty \), for any \( 1 \leq k \leq p \). Then, as \( n \to \infty \)

\[
\sup_{t \in [0,1]} |S_{nk}^{*}(t,0) - S_{nk}(t,0)| = O_{p}(1). \tag{3.2.56}
\]

Proof. Let a weighted empirical process given by

\[
Z_{nk}(t) = \sum_{i=1}^{n} c_{nik} [I\{ \xi_{ni} \leq G_{ni}(t) \} - G_{ni}(t)],
\]

where \( G_{ni} \) are arbitrary d.f.'s on \([0,1]\) and \( \xi_{ni} \) are independent r.v.'s with d.f. \( G_{ni} \), respectively. Then, for \( G_{ni} = U \) and \( \xi_{ni} = F(Y_{i}) \), we have

\[
Z_{nk}(t) = S_{nk}^{*}(t,0) - S_{nk}(t,0) = \sum_{i=1}^{n} c_{nik} [I\{ F(Y_{i}) \leq t \} - t],
\]

and for any \( 0 \leq s, t \leq 1 \),

\[
\text{Cov}\{Z_{nk}(s), Z_{nk}(t)\} = \sum_{i=1}^{n} c_{nik}^{2} \{s \wedge t - st\} = \{s \wedge t - st\}.
\]

Hence, by Theorem 1 of Shorack and Wellner (1986, page 109), we know that \( Z_{nk} \) weakly converges to some \( Z \) on \((D[0,1], \mathcal{D}, \| \cdot \|)\), where \( \| \cdot \| \) is the uniform metric on \( D[0,1] \) and \( \mathcal{D} = \) the \( \sigma \)-field of subsets of \( D[0,1] \) generated by the open balls. Therefore, (3.2.56) follows. \( \Box \)
COROLLARY 3.2.5. For any fixed $k=1, \ldots, p$, assume (A1), (A2) and (B). Then, for any $K>0$, as $n \to \infty$
\[
\sup \{|S^{*}_{nk}(t, u) - S_{nk}(t, 0)|; t \in [0,1], |u| \leq K\} = \mathcal{O}_{p}(1). \tag{3.2.57}
\]

Therefore, as $n \to \infty$
\[
\sup \left\{ \left| \frac{S^{*}_{nk}(t, u)}{\sum_{i=1}^{n} c_{nik}} - t \right|; t \in [0,1], |u| \leq K \right\} \to 0. \tag{3.2.58}
\]

Proof. Since for each $t \in [0,1], |u| \leq K,$
\[
S^{*}_{nk}(t, u) - S_{nk}(t, 0) = [S^{*}_{nk}(t, u) - S_{nk}(t, u)] + [S_{nk}(t, u) - S_{nk}(t, 0)]
\]
\[
= \left\{ [S^{*}_{nk}(t, u) - S_{nk}(t, u)] - [S^{*}_{nk}(t, 0) - S_{nk}(t, 0)] \right\} +
\]
\[
+ [S^{*}_{nk}(t, 0) - S_{nk}(t, 0)] + [S_{nk}(t, u) - S_{nk}(t, 0)],
\]
then, (3.2.57) follows from (3.2.47), Lemma 3.2.4 and Corollary 2.4.3. Furthermore, (3.2.58) follows from (3.2.37). \qed

COROLLARY 3.2.6. Assume (A3) and (B). Then, for any $K>0$ and $1 \leq k \leq p$, as $n \to \infty$
\[
\sup\{S_{nk}^+(t, u) - S_{nk}^+(t, 0); \ t \in [0,1], \ |u| \leq K\} = O_p(1), \quad (3.2.59)
\]

\[
\sup\{S_{nk}^-(t, u) - S_{nk}^-(t, 0); \ t \in [0,1], \ |u| \leq K\} = O_p(1); \quad (3.2.60)
\]

and furthermore, as \( n \to \infty \)

\[
\sup\left\{ \left| \frac{S_{nk}^+(t, u)}{\sum_{i=1}^{n} c_{nik}^+} - t \right|; \ t \in [0,1], \ |u| \leq K \right\} \overset{p}{\to} 0, \tag{3.2.61}
\]

\[
\sup\left\{ \left| \frac{S_{nk}^-(t, u)}{\sum_{i=1}^{n} c_{nik}^-} - t \right|; \ t \in [0,1], \ |u| \leq K \right\} \overset{p}{\to} 0. \tag{3.2.62}
\]

**Proof.** Consider any \( 1 \leq k \leq p \). If \( d_{nk}^+ = 0 \) or \( d_{nk}^- = 0 \), it is the case of Corollary 3.2.5. If \( d_{nk}^+ > 0 \) and \( d_{nk}^- > 0 \), from Corollary 3.2.5, we have, as \( n \to \infty \)

\[
\frac{1}{d_{nk}^+} \sup\{S_{nk}^+(t, u) - S_{nk}^+(t, 0); \ t \in [0,1], \ |u| \leq K\} = O_p(1).
\]

Since \( 0 < d_{nk}^+ < 1 \) and

\[
\frac{S_{nk}^+(t, u)}{\sum_{i=1}^{n} c_{nik}^+} - t = \frac{[S_{nk}^+(t, u) - S_{nk}^+(t, 0)]/d_{nk}^+}{\sum_{i=1}^{n} c_{nik}^+/d_{nk}^+},
\]

and since (A3) implies that

\[
\sum_{i=1}^{n} \frac{c_{nik}^+}{d_{nk}^+} \to \infty, \quad \text{as} \ n \to \infty
\]
(3.2.59) and (3.2.61) follow. The proof of (3.2.60) and (3.2.62) is similar.

\[ \square \]

PROPOSITION 3.2.7. For a fixed \( k=1, \ldots, p \), assume (A1), (A2) and (B). Let \( m: [0,1] \rightarrow \mathbb{R} \) be a function such that \( m \in L^1[0,1] \) and \( m(t)=0 \) outside of an interval \( [a_o,b_o] \), where \( 0<a_o<b_o<1 \), and let a function \( g \) defined on \( \mathbb{R} \) given by

\[
g(x) = \begin{cases} 
F^{-1}(x) & -\infty < x \leq a \\
[x-a+F^{-1}(a)F'(F^{-1}(a))]/F'(F^{-1}(a)) & a \leq x \leq b \\
[x-b+F^{-1}(b)F'(F^{-1}(b))]/F'(F^{-1}(b)) & b \leq x < \infty,
\end{cases}
\] (3.2.63)

where \( 0<a<a_o<b_o<b<1 \). Then, for any \( K>0 \), as \( n \to \infty \)

\[
\sup_{|y| \leq K} \left| \sum_{i=1}^{n} c_{nik} \int_{0}^{t} \left( g(\frac{S^*_{nk}(t,y)}{\sum_{i=1}^{n} c_{nik}}) - g(\frac{S^*_{nk}(t,0)}{\sum_{i=1}^{n} c_{nik}}) \right) m(t) dt - \sum_{i=1}^{n} c_{nik} \xi_{ni}^{T} \int_{0}^{t} m(t) dt \right| \overset{P}{\to} 0.
\] (3.2.64)

Proof. First, we notice that the function \( g \) is continuous and differentiable with a bounded and continuous \( g' \), and that

\[
\sum_{i=1}^{n} c_{nik} \int_{0}^{t} \left( g(\frac{S^*_{nk}(t,y)}{\sum_{i=1}^{n} c_{nik}}) - g(\frac{S^*_{nk}(t,0)}{\sum_{i=1}^{n} c_{nik}}) \right) m(t) dt
\]

\[
= \int_{0}^{t} g'(\xi_{i}) [S^*_{nk}(t,y) - S^*_{nk}(t,0)] m(t) dt
\]
\[
= \int_0^1 [g'(\xi_t) - g'(t)] [S^*_{nk}(t,u) - S^*_{nk}(t,0)] m(t) dt + \\
+ \int_0^1 g'(t) \{[S^*_{nk}(t,u) - S^*_{nk}(t,0)] - \sum_{i=1}^n c_{nik} \xi^T_{ni} u F'(F^{-1}(t))\} m(t) dt + \\
+ \int_0^1 g'(t) \sum_{i=1}^n c_{nik} \xi^T_{ni} u F'(F^{-1}(t)) m(t) dt,
\]

(3.2.65)

where \( \xi_t \) is between \( S^*_{nk}(t,u)/\sum_{i=1}^n c_{nik} \) and \( S^*_{nk}(t,0)/\sum_{i=1}^n c_{nik} \). By Corollary 3.2.2, Corollary 3.2.5 and the uniform continuity of

\[
g'(t) = \frac{1}{F'(F^{-1}(t))}, \quad \text{for } t \in [a, b],
\]

we have, as \( n \to \infty \)

\[
\sup_{|y| \leq K} \left| \int_0^1 [g'(\xi_t) - g'(t)] [S^*_{nk}(t,u) - S^*_{nk}(t,0)] m(t) dt \right| \overset{P}{\to} 0.
\]

(3.2.67)

By the boundedness of \( g' \) on [a, b], \( m \in L^1[0,1] \) and (3.2.66), we have, as \( n \to \infty \)

\[
\sup_{|y| \leq K} \left| \int_0^1 g'(t) \{[S^*_{nk}(t,u) - S^*_{nk}(t,0)] - \sum_{i=1}^n c_{nik} \xi^T_{ni} u F'(F^{-1}(t))\} m(t) dt \right| \overset{P}{\to} 0.
\]

(3.2.68)

Therefore, (3.2.64) follows from (3.2.65) through (3.2.68).

\( \square \)

In the general case of \( \{c_{nik}\} \), we have the following proposition.
PROPOSITION 3.2.8. Under all the assumptions of Proposition 3.2.7, except replacing (A1) and (A2) by (A3), we have that, for any $K > 0$ and $1 \leq k \leq p$, as $n \to \infty$

$$\sup_{|y| \leq K} \left| \sum_{i=1}^{n} c_{nik}^+ \int_{0}^{t} \left[ g \left( \frac{S_{nk}^+(t,y)}{\sum_{i=1}^{n} c_{nik}^+} \right) - g \left( \frac{S_{nk}^+(t,0)}{\sum_{i=1}^{n} c_{nik}^+} \right) \right] m(t) dt - \sum_{i=1}^{n} c_{nik}^+ e_{ni}^+ y \int_{0}^{t} m(t) dt \right| \to 0,$$

(3.2.69)

and

$$\sup_{|y| \leq K} \left| \sum_{i=1}^{n} c_{nik}^- \int_{0}^{t} \left[ g \left( \frac{S_{nk}^-(t,y)}{\sum_{i=1}^{n} c_{nik}^-} \right) - g \left( \frac{S_{nk}^-(t,0)}{\sum_{i=1}^{n} c_{nik}^-} \right) \right] m(t) dt - \sum_{i=1}^{n} c_{nik}^- e_{ni}^- y \int_{0}^{t} m(t) dt \right| \to 0.$$

(3.2.70)

Therefore, as $n \to \infty$

$$\sup_{|y| \leq K} \left| \sum_{i=1}^{n} c_{nik}^+ \int_{0}^{t} \left[ g \left( \frac{S_{nk}^+(t,y)}{\sum_{i=1}^{n} c_{nik}^+} \right) - g \left( \frac{S_{nk}^+(t,0)}{\sum_{i=1}^{n} c_{nik}^+} \right) \right] m(t) dt - \sum_{i=1}^{n} c_{nik}^+ e_{ni}^+ y \int_{0}^{t} m(t) dt \right| \to 0.$$

(3.2.71)

Proof. Consider any $1 \leq k \leq p$. If $d_{nk}^+ = 0$ or $d_{nk}^- = 0$, this is just the case of
Proposition 3.2.7, and (3.2.69) or (3.2.70) is equal to 0. If $d_{nk}^+ > 0$ and $d_{nk}^- > 0$, the proof is the same as Proposition 3.2.7's by using Corollary 3.2.3 and Corollary 3.2.6.

3.3 Asymptotic Normality

LEMMA 3.3.1. Assume (A4), (C2) and (C4), and assume that $a_n \to 0$, as $n \to \infty$. Then,

(1) $M_n(0) \overset{P}\to N_p(0, \sigma^2 Q)$, as $n \to \infty$; \hspace{1cm} (3.3.1)

(2) In addition, assume (C1) and (C3), and assume, for any $K > 0$, we have, as $n \to \infty$

$$\sup_{|u| \leq K} |M_n(u) - M_n(0) + Q_n u \gamma| \overset{P}\to 0,$$ \hspace{1cm} (3.3.2)

then, for $n \geq 1$, $\hat{u}_n$ exists as a solution of the following equations:

$$M_n(u) = 0 \hspace{1cm} \text{(with respect to } u),$$ \hspace{1cm} (3.3.3)

and

$$|Q_n \hat{u}_n - M_n(0)/\gamma| \overset{P}\to 0, \hspace{1cm} \text{as } n \to \infty.$$ \hspace{1cm} (3.3.4)
Therefore, as $n \to \infty$

$$\hat{u}_n = \mathcal{O}_n \left( \hat{\beta}_n - \beta \right) \mathcal{D} \quad N_p(0, \frac{\sigma^2}{\gamma^2} Q^{-1}). \quad (3.3.5)$$

Proof. (1) Recall (3.2.1), we know

$$M_n(\theta) = \sum_{i=1}^{n} \xi_{ni} \psi(Y_i),$$

where $Y_i$ are i.i.d.r.v's with d.f. $F$. It suffices to show that, for an arbitrary vector

$$\alpha = (a_1, \ldots, a_p)^\tau \in \mathbb{R}^p,$$

$$\alpha^\tau M_n(\theta) \mathcal{D} N(0, \sigma^2_\alpha^\tau Q_\alpha), \quad \text{as } n \to \infty. \quad (3.3.6)$$

Since

$$\alpha^\tau M_n(\theta) = \sum_{i=1}^{n} \alpha^\tau \xi_{ni} \psi(Y_i),$$

letting

$$X_{ni} = \frac{\alpha^\tau \xi_{ni} \psi(Y_i)}{\sigma \sqrt{\alpha^\tau Q_\alpha \alpha}}, \quad i=1, \ldots, n$$

then, $X_{n1}, \ldots, X_{nn}$ are independent r.v.'s, and satisfy

$$E\{X_{ni}\} = 0,$$
\[
\frac{\sum}{i=1} \text{Var}(X_{ni}) = \frac{\sum}{i=1} \left( \frac{(a^T \varepsilon_{ni})^2 \sigma^2}{\vartheta_n^T Q_n \vartheta_n} \right) = \frac{\sum}{i=1} \left( \frac{a^T \varepsilon_{ni} \varepsilon_{ni}^T a}{\vartheta_n^T Q_n \vartheta_n} \right) = 1.
\]

Also, let \( F_{ni} \) be the d.f. of \( X_{ni} \); then, for any \( \eta > 0 \)

\[
\frac{\sum}{i=1} \int |x| > \eta \ x^2 \ dF_{ni}(x) = \frac{\sum}{i=1} E\{X_{ni} I(|X_{ni}| > \eta)\}
\]

\[
= \frac{\sum}{i=1} E\left\{ \frac{(a^T \varepsilon_{ni})^2 \psi^2(Y_i)}{\sigma^2 a^T Q_n \vartheta_n} I\{|\psi(Y_i)| > \eta \frac{a^T Q_n \vartheta_n}{\vartheta_n^T \varepsilon_{ni}}\} \right\}
\]

\[
= \frac{\sum}{i=1} \frac{(a^T \varepsilon_{ni})^2}{\sigma^2 a^T Q_n \vartheta_n} \int \psi^2(y) I\{|\psi(y)| > \eta \frac{a^T Q_n \vartheta_n}{\vartheta_n^T \varepsilon_{ni}}\} dF(y)
\]

\[
\leq \frac{1}{\sigma^2} \int \psi^2(y) I\{|\psi(y)| > \eta \frac{\sigma M}{|\vartheta_n| a_n}\} dF(y),
\]

where \( M > 0 \) is a constant. Hence, from (C4) and the assumption: \( a_n \to 0 \), as \( n \to \infty \), we have

\[
\frac{\sum}{i=1} \int |x| > \eta \ x^2 \ dF_{ni}(x) \to 0, \quad \text{as } n \to \infty.
\]

Therefore, by Theorem 7.2.1 of Chung (1974), we have

\[
\frac{\sum}{i=1} X_{ni} = \frac{a^T M_n(0)}{\sigma \sqrt{a^T Q_n \vartheta_n}} \mathcal{D} N(0,1), \quad \text{as } n \to \infty.
\]

Hence, (3.3.6) follows from (A4).

(2) (C1) implies that \( \rho \) is convex and that there exist \( \alpha \) and \( \beta \) such that
\[ \psi(\alpha) < 0, \quad \text{and} \quad \psi(\beta) > 0; \quad \alpha < \beta. \] (3.3.7)

So, for \( x < \alpha \) and \( y > \beta \), we have

\[ \frac{\rho(x) - \rho(\alpha)}{x - \alpha} \leq \psi(\alpha) < 0 \quad \text{and} \quad \frac{\rho(y) - \rho(\beta)}{y - \beta} \geq \psi(\beta) > 0. \]

Therefore,

\[ \lim_{x \to \pm \infty} \rho(x) = +\infty, \] (3.3.8)

and \( \rho \) has a finite lower bound. Let \( A_n \) be a \( p \times p \) non-singular matrix such that

\[ A_n^T Q_n A_n \] is a diagonal matrix which (by (A4)) only has positive eigenvalues for large enough \( n \). Consider a function \( \Phi_n \) defined as below:

\[ \Phi_n(y) = \sum_{i=1}^{n} \rho(Y_i - \varepsilon_n^T A_n y), \quad n \geq 1, \quad \forall y \in \mathbb{R}^p. \] (3.3.9)

Then, for any \( y \in \mathbb{R}^p \), we have, as \( \|y\| \to \infty \)

\[ \sum_{i=1}^{n} (\varepsilon_n^T A_n y)^2 = \sum_{i=1}^{n} y^T A_n^T \varepsilon_n \varepsilon_n^T A_n y = y^T A_n^T (\sum_{i=1}^{n} \varepsilon_n \varepsilon_n^T) A_n y \]

\[ = y^T A_n^T Q_n A_n y \geq \lambda_1(Q_n) \|y\|^2 \to \infty, \]

where \( \lambda_1(A) \) denotes the smallest eigenvalue of the matrix \( A \). Hence, by (3.3.8), for any large enough \( n \),
\[ \Phi_n(y) \to \infty, \quad \text{as} \quad \|y\| \to \infty. \tag{3.3.10} \]

Since \( \rho \) is convex, \( \Phi_n \) is also a convex function for \( n \geq 1 \). Therefore, \( \Phi_n \) has at least one global minimum, i.e., by the convexity of \( \rho \), (3.1.4), equivalent to (3.1.5), has a finite solution. Therefore, from the equivalent relation between (3.1.5) and (3.2.7), (3.3.3) has a solution for large enough \( n \).

From (3.3.1), for any \( \epsilon > 0 \), there exist \( K_0 > 0 \) and \( N \) such that

\[ P\{|M_n(\theta)|>K_0\} < \epsilon, \quad \text{for } n \geq N. \tag{3.3.11} \]

For arbitrary \( \rho > 0 \) and \( K > 0 \), since (3.3.3) has a solution \( \hat{u}_n \) for \( n \geq 1 \),

\[ P\{|Q_n \hat{u}_n \gamma - M_n(\theta)|>\rho\} \]

\[ = P\{|Q_n \hat{u}_n \gamma - M_n(\theta)|>\rho, \|\hat{u}_n\| \leq K\} + P\{|Q_n \hat{u}_n \gamma - M_n(\theta)|>\rho, \|\hat{u}_n\| > K\} \]

\[ \leq P\left\{ \sup_{|u| \leq K} \left| M_n(u) - M_n(\theta) + Q_n u \gamma \right| \geq \rho \right\} + P\{\|\hat{u}_n\| > K\}. \tag{3.3.12} \]

Consider a function given by

\[ h(y, B) = y^T M_n(uB) \]

\[ = \sum_{i=1}^{n} y^T \xi_{ni} \psi(Y_i - \xi_{ni}^T y B), \quad y \in \mathbb{R}^p, B \in (0, +\infty). \tag{3.3.13} \]

Since \( \psi \) is nondecreasing, \( h(y, B) \) is nonincreasing in \( B \) for any \( y \in \mathbb{R}^p \). Hence, we
claim: $\|\hat{y}_n\|>K$ implies that

$$h(e_n, K) \geq 0,$$

(3.3.14)

where $e_n = \hat{y}_n/\|\hat{y}_n\| \in \mathbb{R}^P$. Otherwise, $h(e_n, K) < 0$ implies

$$h(e_n, \|\hat{y}_n\|) \leq h(e_n, K) < 0,$$

(3.3.15)

which, by the definition of $\hat{y}_n$ and (3.3.13), contradicts

$$h(e_n, \|\hat{y}_n\|) = e_n^T M_n(\hat{y}_n) = 0.$$

Therefore, by (3.3.11), for $n \geq N$,

$$P\{|\|\hat{y}_n\|>K\| \leq P\{e_n^T M_n(e_n K) \geq 0\}$$

$$= P\{e_n^T [M_n(e_n K) - M_n(0) + Q_n e_n K \gamma] \geq -e_n^T M_n(0) + e_n^T Q_n e_n K \gamma\}$$

$$\leq P\{e_n^T [M_n(e_n K) - M_n(0) + Q_n e_n K \gamma] \geq -pK_0 + \lambda_1(K_n K \gamma) + \epsilon\}.$$

Therefore, by (A4), when $n$ large enough, there exists a positive number $\lambda$ such that $0 < \lambda < \lambda_1(Q)$ and

$$P\{|\|\hat{y}_n\|>K\| \leq P\{e_n^T [M_n(e_n K) - M_n(0) + Q_n e_n K \gamma] \geq -pK_0 + \lambda K \gamma + \epsilon\}.$$
Choose $K$ such that $K > \frac{pK_0}{\lambda \gamma}$, then $\beta_1 = -pK_0 + \lambda K \gamma > 0$, and by (3.3.2), we have

\[ P\{\|\hat{u}_n\| > K\} \leq P\{\sup_{|u| \leq K} |M_n(u) - M_n(0) + Q_n u \gamma| \geq \beta_1\} + \epsilon \rightarrow \epsilon. \tag{3.3.16} \]

Therefore, (3.3.4) follows from (3.3.12), (3.3.2) and (3.3.16).

THEOREM 3.3.2. Let $\psi$ be a bounded function satisfying Lipschitz-condition of order $\alpha > \frac{1}{2}$, and $F$ be an absolutely continuous d.f. with a bounded and continuous derivative $F'$. Assume (C1) through (C3) and assume (A4), and

\[ a_n \rightarrow 0, \quad \text{as } n \rightarrow \infty. \tag{3.3.17} \]

Then, for any $K > 0$ and $1 \leq k \leq p$, as $n \rightarrow \infty$

\[ \sup_{|u| \leq K} |\int \psi(x) dS_{nk} (F(x), u) + \sum_{i=1}^{n} \epsilon_{ni} \epsilon_{ni}^r u \gamma | \rightarrow 0, \tag{3.3.18} \]

\[ \sup_{|u| \leq K} |M_{nk}(u) - M_{nk}(0) - \int \psi(x) dS_{nk} (F(x), u) | \rightarrow 0. \tag{3.3.19} \]

Therefore, as $n \rightarrow \infty$

\[ \sup_{|u| \leq K} |M_n(u) - M_n(0) + Q_n u \gamma | \rightarrow 0. \tag{3.3.20} \]

and furthermore,
\[ C_n^0 \left( \hat{\theta}_n - \theta \right) \overset{D}{\sim} N_p(0, \frac{\sigma^2}{\gamma} Q^{-1}). \quad (3.3.21) \]

Proof. Since, by Fubini Theorem,

\[
\int_{-\infty}^{\infty} \psi(x) dF(x + \zeta_{ni}^T u) = \int_{-\infty}^{\infty} \psi(y - \zeta_{ni}^T u) dF(y)
\]

\[
= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x < y - \zeta_{ni}^T u) d\psi(x) dF(y) + \psi(-\infty)
\]

\[
= \int_{-\infty}^{\infty} [1 - F(x + \zeta_{ni}^T u)] d\psi(x) + \psi(-\infty)
\]

\[
= \psi(\infty) - \int_{-\infty}^{\infty} F(x + \zeta_{ni}^T u) d\psi(x),
\]

and, by (C2), we have

\[
\psi(\infty) = \int_{-\infty}^{\infty} F(x) d\psi(x),
\]

so, we have

\[
\int_{-\infty}^{\infty} \psi(x - \zeta_{ni}^T u) dF(x) = \int_{-\infty}^{\infty} [F(x) - F(x + \zeta_{ni}^T u)] d\psi(x). \quad (3.3.22)
\]

Therefore,

\[
\left| \int \psi(x) dS_{nk}(F(x), u) + \sum_{i=1}^{\mathcal{N}} c_{nik} \zeta_{ni}^T u \gamma \right|
\]
\[
\begin{align*}
&= \left| \sum_{i=1}^{P} c_{ni} \int \psi(x) dF(x + \varepsilon_{ni}^T u) + \sum_{i=1}^{P} c_{ni} \varepsilon_{ni}^T u \int \psi'(x) dF(x) \right| \\
&= \left| \sum_{i=1}^{P} c_{ni} \left\{ \left[ F(x) - F(x + \varepsilon_{ni}^T u) \right] d\psi(x) + \varepsilon_{ni}^T u \int F'(x) d\psi(x) \right\} \right| \\
&= \left| \sum_{i=1}^{P} c_{ni} \varepsilon_{ni}^T u \left[ F'(\xi) - F'(x) \right] d\psi(x) \right| \\
&\leq \sum_{i=1}^{P} \sum_{j=1}^{P} |c_{ni}| |c_{nj}| |u| \int \left| F'(\xi) - F'(x) \right| d\psi(x) \\
&\leq \sum_{j=1}^{P} \sqrt{\sum_{i=1}^{P} c_{ni}^2} \sqrt{\sum_{i=1}^{P} c_{nj}^2} |u| \int \sup_{i,j} \left| F'(\xi) - F'(x) \right| d\psi(x) \\
&= p|u| \int \sup_{i,j} \left| F'(\xi) - F'(x) \right| d\psi(x)
\end{align*}
\]

where \( \xi \) is between \( x \) and \( x + \varepsilon_{ni}^T u \). Hence, using DCT, (3.3.18) follows from (3.3.17) and the boundedness and continuity of \( F' \).

It is easy to see that, for each \( u \in \mathbb{R}^P \), \( \mathbb{E} \{ M_{nk}(u) \} = \int \psi(x) dS_{nk}(F(x), u) \) and that \( \mathbb{E} \{ M_{nk}(0) \} = 0 \). Then, (3.3.19) follows from showing: for any \( K > 0 \),

\[
\sup_{|u| \leq K} |D_{nk}(u)| \xrightarrow{P} 0, \quad \text{as } n \to \infty. \tag{3.3.23}
\]

where

\[
D_{nk}(u) = [M_{nk}(u) - M_{nk}(0)] - \mathbb{E}[M_{nk}(u) - M_{nk}(0)], \quad u \in \mathbb{R}^P.
\]

Let

\[
I_r = \text{Diag}(e_1, \ldots, e_p), \quad 1 \leq r \leq 2^P \tag{3.3.24}
\]
where \((e_1, \cdots, e_p)\) takes on the \(2^p\) possible realizations on the vertices of \([-1,1]^p\), i.e., \(e_j\) is \(-1\) or \(1\), for \(1 \leq j \leq p\), and let, for \(y \in \mathbb{E}_p\),

\[
B_{nkr}(y) = \left[M_{nk}(IrK) - M_{nk}(0)\right] - E[M_{nk}(IrK) - M_{nk}(0)],
\]

then,

\[
\sup_{|y| \leq K} |D_{nk}(y)| = \max_{1 \leq r \leq 2^p} \left\{ \sup_{y \in \mathbb{E}_p} |B_{nkr}(y)| \right\}.
\]

Hence, it suffices to show that, for any \(1 \leq r \leq 2^p\), as \(n \to \infty\)

\[
\sup_{y \in \mathbb{E}_p} |B_{nkr}(y)| \xrightarrow{p} 0.
\]

We notice that \(E(B_{nkr}(y)) = 0\) and \(B_{nkr}(0) = 0\). Let \(i_1 \in \mathbb{L}_p(m)\), \(0 \leq i < (m+1)^p\), where \(m\) is an arbitrary positive integer and \(j_i\) satisfies that: (i) \(j_0 = 0\); (ii) \(|i_i| = i_{i-1}| = \frac{1}{m}\), for \(1 \leq i < (m+1)^p\). Denote

\[
\xi_i = B_{nkr}(i_i) - B_{nkr}(i_{i-1}), \quad 1 \leq i < (m+1)^p
\]

\[
S_q = \sum_{i=1}^q \xi_i = B_{nkr}(j_q), \quad 1 \leq q < (m+1)^p.
\]

Since \(\psi\) satisfies Lipschitz-condition with order \(\alpha\), then, for any \(y, \gamma \in \mathbb{E}_p\),

\[
\text{Var}(B_{nkr}(y) - B_{nkr}(\gamma)) = \sum_{i=1}^n c_{ni}^2 \text{Var}\{\psi(Y_1 - c_{ni}^T i_r K) \psi(Y_1 - c_{ni}^T i_r K)\}
\]

\[
\leq \sum_{i=1}^n c_{ni}^2 \text{MK} |\xi_{ni}^T (y - \gamma)|^{2\alpha} \leq \text{pMK} \sigma_n^{2\alpha} |y - \gamma|^{2\alpha}
\]
where $M$ is a constant. So, for any $\lambda > 0$ and $i \leq j$,

$$P(|S_j - S_i| \geq \lambda) = P(|B_{nkr}(j) - B_{nkr}(i)| \geq \lambda)$$

$$\leq \frac{\text{Var}\{B_{nkr}(j) - B_{nkr}(i)\}}{\lambda^2} \leq \frac{pMK}{\lambda^2} \frac{a_n^{2\alpha}}{|j - i|^{2\alpha}}$$

$$\leq \frac{pMK}{\lambda^2} \left\{ \frac{j - i}{m} \right\}^{2\alpha} = \frac{pMK}{\lambda^2} \left\{ \sum_{1 \leq l \leq j} u_l \right\}^{2\alpha}$$

where $u_l = \frac{a_n}{m}$. Since $2\alpha > 1$, by Theorem 12.2 of Billingsley (1968, page 94), we have, for any $\rho > 0$,

$$P\left\{ \max_{1 \leq q < (m+1)} |S_q| > \rho \right\} = P\left\{ \max_{1 \leq i < (m+1)} \rho |B_{nkr}(i)| > \rho \right\}$$

$$\leq \frac{M_1}{\rho^2} \left\{ \sum_{i=1}^{m} \frac{a_n}{m} \right\}^{2\alpha} = \frac{M_1}{\rho^2} (a_n)^{2\alpha}$$

where $M_1$ is a constant. Since $B_{nkr}(u)$ is continuous in $u$, letting $m \to \infty$, we have

$$P\left\{ \sup_{u \in E_p} |B_{nkr}(u)| \geq \rho \right\} \leq \frac{M_1}{\rho^2} (a_n)^{2\alpha}.$$ 

Therefore, (3.3.27) follows from (3.3.17). \qed

THEOREM 3.3.3. For a fixed $k = 1, \ldots, p$, assume (A1), (A2), (B), (C1), (C2) and (C5). Then, for any $K > 0$, as $n \to \infty$
\[ \sup_{|u| \leq K} \left| M_{nk}(u) - M_{nk}(0) + \sum_{i=1}^{n} c_{nik} e_{ni}^T u \gamma_0 \right| \leq 0, \quad (3.3.30) \]

where \( \gamma_0 = \int F' d\psi > 0. \)

Proof. Let \( M: [0,1] \to \mathbb{R} \) be a function defined by \( M(t) = \psi(F^{-1}(t)), 0 \leq t \leq 1. \)

From the assumptions on \( \psi \) and \( F \), we have that \( M \) is nondecreasing, right continuous and bounded. Consider a functional \( \tau: D[0,1] \to \mathbb{R} \) defined by

\[ \tau(G) = \int_0^1 h(G(t)) e^t dM(t), \quad G \in D[0,1] \quad (3.3.31) \]

where \( h \) is defined by

\[ h(x) = \begin{cases} 
- e^{-x} & \text{if } x \geq 0 \\
(3.3.32)
x - 1 & \text{if } x < 0.
\end{cases} \]

Then, \( \tau \) can be expressed as a composition of the following Hadamard differentiable transformations:

\[ \gamma_1: D[0,1] \to L^1[0,1], \quad \gamma_1(S) = h \circ S; \]

\[ \gamma_2: L^1[0,1] \cap D[0,1] \to \mathbb{R}, \quad \gamma_2(S) = \int_0^1 S(t) e^t dM(t); \]
\( \tau(G) = \gamma_2(\gamma_1(G)) \).

Note that \( \gamma_1 \) is Hadamard differentiable at \( U \) by Proposition 6.1.2 of Fernholz (1983) (which is also true for \( \gamma: D[0,1] \rightarrow L^q[0,1] \)), because \( h \) is differentiable everywhere with a bounded and continuous derivative, and that \( \gamma_2 \) is linear and continuous, thus Fréchet differentiable. Since the range of \( \gamma_1 \) is contained by \( D[0,1] \), from Proposition 3.1.2 of Fernholz (1983), \( \tau \) is Hadamard differentiable at \( U \) with Hadamard derivative

\[
\tau'_U(G) = \int_0^1 h'(t) G(t) e^t dM(t), \quad G \in D[0,1]. \tag{3.3.33}
\]

By Theorem 2.3.1, we have, for any \( K > 0 \), as \( n \to \infty \)

\[
\sup_{|u| \leq K} \left| \sum_{i=1}^{n} c_{nik} \left\{ \left( \frac{S_{nk}^*(\cdot, y)}{\sum_{i=1}^{n} c_{nik}} - \tau(U(\cdot)) \right) - \tau'_U(S_{nk}^*(\cdot, y) - U(\cdot)) \sum_{i=1}^{n} c_{nik} \right\} \right| \to 0.
\]

Therefore, as \( n \to \infty \)

\[
\sup_{|u| \leq K} \left| \sum_{i=1}^{n} c_{nik} \left\{ \left( \frac{S_{nk}^*(\cdot, y)}{\sum_{i=1}^{n} c_{nik}} - \tau\left( \frac{S_{nk}^*(\cdot, 0)}{\sum_{i=1}^{n} c_{nik}} \right) \right) - \tau'_U(S_{nk}^*(\cdot, y) - S_{nk}^*(\cdot, 0)) \right\} \right| \to 0. \tag{3.3.34}
\]

Since, \( h'(t) = e^t \), for \( 0 \leq t \leq 1 \), then, for each \( y \),
\[ \tau_u(S_{nk}^*(\cdot, u)) = \int_0^1 h'(t)S_{nk}^*(t, u) e^t dM(t) = \sum_{i=1}^n c_{nik} \int_{F(\gamma_1 - \varepsilon_{ni}^T u)}^\infty d\psi(x) = \psi(\infty) \sum_{i=1}^n c_{nik} - M_{nk}(u). \] 

(3.3.35)

By the similar proof of Proposition 3.2.7, we have

\[ \sup_{|u| \leq K} \left| \sum_{i=1}^n c_{nik} \left\{ \tau\left( \frac{S_{nk}^*(\cdot, u)}{\sum_{i=1}^n c_{nik}} \right) - \tau\left( \frac{S_{nk}^*(\cdot, 0)}{\sum_{i=1}^n c_{nik}} \right) \right\} - \sum_{i=1}^n c_{nik} c_{ni}^T u \int_0^1 F'(F^{-1}(t)) d\psi(F^{-1}(t)) \right| \xrightarrow{P} 0. \] 

(3.3.36)

Hence, (3.3.30) follows from (3.3.34) through (3.3.36).

In the general case of \( \{c_{nik}\} \), \( 1 \leq i \leq n, 1 \leq k \leq p \), we have the following theorem.

**THEOREM 3.3.4.** Under all the assumptions of Theorem 3.3.3, except replacing (A1) and (A2) by (A3) and (A4), we have that, for any \( K > 0 \) and any \( 1 \leq k \leq p \), as \( n \to \infty \)

\[ \sup_{|u| \leq K} \left| M_{nk}(u) - M_{nk}(0) + \sum_{i=1}^n c_{nik} c_{ni}^T u \gamma_0 \right| \xrightarrow{P} 0. \] 

(3.3.37)

Therefore, as \( n \to \infty \)

\[ \sup_{|u| \leq K} \left| M_n(u) - M_n(0) + Q_n u \gamma_0 \right| \xrightarrow{P} 0, \] 

(3.3.38)
and furthermore,

$$C_n^0 ( \hat{\beta}_n - \beta ) \overset{d}{\rightarrow} N_p (0, \frac{\sigma^2}{\gamma_0} Q^{-1}). \quad (3.3.39)$$

Proof. Consider any $1 \leq k \leq p$. If $d_{nk}^+ = 0$ or $d_{nk}^- = 0$, (3.3.37) is the case of Theorem 3.3.3. If $d_{nk}^+ > 0$ and $d_{nk}^- > 0$, using the functional $\tau$ given by (3.3.31) and Theorem 2.3.2, we have, for any $K > 0$ and $1 \leq k \leq p$, as $n \to \infty$

$$\sup_{|y| \leq K} \left| \sum_{i=1}^{n} c_{nik}^+ \frac{S_{nk}^+ (\cdot, y)}{\sum_{i=1}^{n} c_{nik}^+} - \sum_{i=1}^{n} c_{nik}^- \frac{S_{nk}^- (\cdot, y)}{\sum_{i=1}^{n} c_{nik}^-} - \tau(U) \sum_{i=1}^{n} c_{nik}^+ - \tau(U) (S_{nk}^+ (\cdot, y) - U (\cdot) \sum_{i=1}^{n} c_{nik}^-) \right| \overset{P}{\to} 0. \quad (3.3.40)$$

So, by (3.3.35), we have, as $n \to \infty$

$$\sup_{|y| \leq K} \left| \sum_{i=1}^{n} c_{nik}^+ \left\{ \tau \left( \frac{S_{nk}^+ (\cdot, y)}{\sum_{i=1}^{n} c_{nik}^+} \right) - \tau \left( \frac{S_{nk}^+ (\cdot, y)}{\sum_{i=1}^{n} c_{nik}^+} \right) \right\} - \sum_{i=1}^{n} c_{nik}^- \left\{ \tau \left( \frac{S_{nk}^- (\cdot, y)}{\sum_{i=1}^{n} c_{nik}^-} \right) - \tau \left( \frac{S_{nk}^- (\cdot, y)}{\sum_{i=1}^{n} c_{nik}^-} \right) \right\} + \frac{M_{nk} (y) - M_{nk} (0)}{\sum_{i=1}^{n} c_{nik}^-} \right| \overset{P}{\to} 0. \quad (3.3.41)$$

By the similar proof of Proposition 3.2.8, we have

$$\sup_{|y| \leq K} \left| \sum_{i=1}^{n} c_{nik}^+ \left\{ \tau \left( \frac{S_{nk}^+ (\cdot, y)}{\sum_{i=1}^{n} c_{nik}^+} \right) - \tau \left( \frac{S_{nk}^+ (\cdot, y)}{\sum_{i=1}^{n} c_{nik}^+} \right) \right\} - \sum_{i=1}^{n} c_{nik}^- \left\{ \tau \left( \frac{S_{nk}^- (\cdot, y)}{\sum_{i=1}^{n} c_{nik}^-} \right) - \tau \left( \frac{S_{nk}^- (\cdot, y)}{\sum_{i=1}^{n} c_{nik}^-} \right) \right\} \right|$$
\[- \sum_{i=1}^{n} c_{nik} \{ \tau \left( \frac{S_{nk}^*(\cdot, \mathbf{u})}{\sum_{i=1}^{n} c_{nik}} \right) - \tau \left( \frac{S_{nk}^*(\cdot, 0)}{\sum_{i=1}^{n} c_{nik}} \right) \} \sum_{i=1}^{n} c_{nik} c_{ni}^T \mathbf{y}_0 \] \quad \mathbb{P} \rightarrow 0.

(3.3.42)

Therefore, (3.3.37) follows from (3.3.41) and (3.3.42).

\[ \square \]

**Remark.** Comparing Theorem 3.3.4 with Theorem 3.3.2, we can see that we require weaker conditions on \( \psi \), but a slightly stronger conditions on \( F \) in Theorem 3.3.4 by incorporating Hadamard differentiability through the uniform asymptotic linearity of the M-estimators of regression.

3.4 Consistency

Using the Hadamard differentiability approach, which is used in last section for the asymptotic normality, we have the following theorem for the weak consistency of the M-estimator of regression.

**THEOREM 3.4.1.** In addition to (A3), (B) and (C1) through (C4), assume that \( \inf_{n \geq 1} \{ \lambda_1(Q_n) \} > 0 \) and \( \psi' = 0 \) outside of a finite interval. Then,

\[ |\hat{\beta}_n - \beta| \xrightarrow{\mathbb{P}} 0, \quad \text{as } n \to \infty. \]

(3.4.1)
Proof. Let \( m: [0,1] \rightarrow \mathbb{R} \) be a function defined by \( m(t) = \psi'(F^{-1}(t)), \quad 0 \leq t \leq 1 \). From the assumptions on \( \psi \) and \( F \), we have that \( m \in L^1[0,1] \) and \( m(t) = 0 \) outside of an interval \([a_0, b_0]\), where \( 0 < a_0 < b_0 < 1 \). For \( 0 < a < a_0 < b_0 < b < 1 \), consider the function \( g \) given by (3.2.63) and consider a functional \( \tau: D[0,1] \rightarrow \mathbb{R} \) defined by

\[
\tau(G) = \int_0^1 g(G(t)) \, m(t) \, dt, \quad G \in D[0,1]. \tag{3.4.2}
\]

Then, \( \tau \) can be expressed as a composition of the following Hadamard differentiable transformations:

\[
\gamma_1: D[0,1] \rightarrow L^1[0,1], \quad \gamma_1(S) = g \circ S;
\]

\[
\gamma_2: L^1[0,1] \cap D[0,1] \rightarrow \mathbb{R}, \quad \gamma_2(S) = \int_0^1 S(t) \, m(t) \, dt;
\]

and

\[
\tau(G) = \gamma_2(\gamma_1(G)).
\]

Note that \( \gamma_1 \) is Hadamard differentiable at \( U \) by Proposition 6.1.2 of Fernholz (1983) (which is also true for \( \gamma: D[0,1] \rightarrow L^q[0,1] \)), because \( g \) is differentiable everywhere with a bounded derivative, and that \( \gamma_2 \) is linear and continuous, thus Fréchet differentiable. Since the range of \( \gamma_1 \) is contained by \( D[0,1] \), from Proposition 3.1.2 of Fernholz (1983), \( \tau \) is Hadamard differentiable at \( U \) with Hadamard derivative

\[
\tau'(G) = \int_0^1 g'(t) \, G(t) \, m(t) \, dt, \quad G \in D[0,1]. \tag{3.4.3}
\]
So, by Theorem 2.3.2, we have, for any $K > 0$, as $n \to \infty$

\[ \sup_{|y| \leq K} \left| \sum_{i=1}^{n} c_{nik}^{+} \tau \left( \frac{S_{nk}^{+}(\cdot,y)}{\sum_{i=1}^{n} c_{nik}^{+}} \right) - \sum_{i=1}^{n} c_{nik}^{-} \tau \left( \frac{S_{nk}^{-}(\cdot,y)}{\sum_{i=1}^{n} c_{nik}^{-}} \right) - \tau(U) \sum_{i=1}^{n} c_{nik} \right| \to 0. \]  

(3.4.4)

Since, $g'(t) = 1/F'(F^{-1}(t))$, for $a \leq t \leq b$, then, for each $y$

\[ \tau(U) = \int_{0}^{1} g'(t)S_{nk}^{+}(t,y)m(t)dt = \sum_{i=1}^{n} c_{nik} \int_{\gamma_{i}-\xi_{ni}^{+}y}^{\gamma_{i}^{+}} \psi'(x)dx. \]  

(3.4.5)

Therefore, by (3.4.4) and (3.4.5), as $n \to \infty$

\[ \sup_{|y| \leq K} \left| \sum_{i=1}^{n} c_{nik}^{+} \left\{ \tau \left( \frac{S_{nk}^{+}(\cdot,y)}{\sum_{i=1}^{n} c_{nik}^{+}} \right) - \tau \left( \frac{S_{nk}^{+}(\cdot,0)}{\sum_{i=1}^{n} c_{nik}^{+}} \right) \right\} - \sum_{i=1}^{n} c_{nik}^{-} \left\{ \tau \left( \frac{S_{nk}^{-}(\cdot,y)}{\sum_{i=1}^{n} c_{nik}^{-}} \right) - \tau \left( \frac{S_{nk}^{-}(\cdot,0)}{\sum_{i=1}^{n} c_{nik}^{-}} \right) \right\} + \sum_{i=1}^{n} c_{nik} \int_{\gamma_{i}-\xi_{ni}^{+}y}^{\gamma_{i}^{+}} \psi'(y)dy \right| \to 0. \]  

(3.4.6)

Hence, by Proposition 3.2.8 and $\gamma = \int \psi'dF = \int_{0}^{1} m(t)dt$, we have that, for any $1 \leq k \leq p$, as $n \to \infty$
\[ \sup_{|u| \leq K} \left| \sum_{i=1}^{\mathcal{B}} c_{n,\gamma} \psi(y) \right| \leq K \infty \leq K \infty \] 

Therefore, as \( n \to \infty \)

\[ \sup_{|u| \leq K} \left| Q_{\gamma} y + \sum_{i=1}^{\mathcal{B}} c_{n,\gamma} \psi(y) \right| \leq K \infty \] 

Denote

\[ \phi_n(\varepsilon, B) = \sum_{i=1}^{\mathcal{B}} \varepsilon^{r_{n,\gamma}} \int_{Y_i} \psi(y) dy, \] 

where \( B \in (0, +\infty) \) and \( \varepsilon \in \mathbb{R}^P \) satisfying \( ||\varepsilon|| = 1 \). Since \( \psi \) is nondecreasing, then, \( \phi_n \) is nonincreasing in \( B \) and we have, for \( B > 0, \Delta \in \mathbb{R} \),

\[ 0 \leq \Delta \int_{Y_i - \Delta B} \psi'(y) dy \leq \Delta [\psi(Y_i) - \psi(Y_i - \Delta B)]. \] 

Hence, for \( B > 0 \),

\[ 0 \leq -\phi_n(\varepsilon, B) = -\sum_{i=1}^{\mathcal{B}} \varepsilon^{r_{n,\gamma}} \int_{Y_i} \psi'(y) dy \leq \varepsilon^{T} [M_n(0) - M_n(\varepsilon B)]. \]

Since, for any \( \rho > 0 \), we have

\[ P\{|\hat{\theta}_n - \theta| > \rho\} = P\{|(C_n^{\gamma})^{-1} \hat{\gamma}_n| > \rho\} \leq P\{|\tilde{\gamma}_n| > \frac{\rho}{||C_n^{\gamma}||_{1}}\}, \]
and since (A3) implies that, as \( n \to \infty \)

\[
\frac{1}{|C_n^0|^{-1} I} \to +\infty,
\]

it suffices to show that, for large enough \( K > 0 \),

\[
P\{\|\hat{u}_n\| > K\} \to 0, \quad \text{as } n \to \infty. \tag{3.4.13}
\]

Since, for arbitrary \( \epsilon > 0, 1 \leq k \leq p \) and \( n \geq 1 \), there exists \( K_0 > 0 \) such that

\[
P\{|M_{nk}(0)| > K_0\} \leq \frac{\text{Var}\{M_{nk}(0)\}}{K_0^2} = \frac{\sigma^2}{K_0^2} < \epsilon,
\]

therefore, for \( n \geq 1 \),

\[
P\{|M_n(0)| > K_0\} < \epsilon. \tag{3.4.14}
\]

Using the similar argument referring to (3.3.16) in the proof of (2) of Lemma 3.3.1, by (3.4.11) and (3.4.14), we have, for large \( n \)

\[
P\{\|\hat{u}_n\| > K\} \leq P\{e_n^\top M_n(e_n K) \geq 0\}
\]

\[
= P\{e_n^\top [M_n(e_n K) - M_n(0) + Q_n e_n K] \geq -e_n^\top M_n(0) + e_n^\top Q_n e_n K\}
\]

\[
\leq P\{[\phi_n(e_n, K) + e_n^\top Q_n e_n K] \geq -e_n^\top M_n(0) + e_n^\top Q_n e_n K\}
\]
\begin{equation}
\begin{aligned}
\leq & \mathbb{P}\left\{ \sup_{|y| \leq K} \left| Q_n y + \sum_{i=1}^{n} \epsilon_{ni} \psi'(y) dy \right| \geq -\epsilon_n^T M_n(\theta) + \epsilon_n^T Q_n e_n K \gamma \right\} \\
\leq & \mathbb{P}\left\{ \sup_{|y| \leq K} \left| Q_n y + \sum_{i=1}^{n} \epsilon_{ni} \psi'(y) dy \right| \geq \beta_1 \right\} + \epsilon, \quad (3.4.15)
\end{aligned}
\end{equation}

where \( \beta_1 = -pK_0 + K_7\left(\inf\{\lambda_1(Q_n)\}\right) > 0 \), for \( K \) large enough. Therefore, (3.4.13) follows from (3.4.8) and (3.4.15). \( \square \)

However, the Hadamard differentiability approach may not provide a way for the study of the strong consistency, which naturally requires stronger conditions on \( \psi, F \) and \( \{\epsilon_{ni}\}, \, 1 \leq i \leq n \). In this context, a different approach is also considered for the weak consistency as below, which, later on, will lead to the conditions for the strong consistency in Theorem 3.4.4 and Theorem 3.4.5.

**THEOREM 3.4.2.** Assume that, in addition to (C1) and (C2), \( \psi, F \) and \( \mathcal{C}_n \) satisfy the following conditions:

1. (N1) \( \psi \) is bounded, right continuous and piecewise differentiable with a bounded derivative \( \psi' \)
2. (N2) \( F \) is a d.f. with a bounded and continuous derivative \( F' \)
3. (N3) \( \gamma_0 = \int F'd\psi > 0 \)
4. (N4) There exists a constant \( M \) such that \( \max_{1 \leq i \leq n} \|\xi_i\| \leq M < \infty \)
5. (N5) Let \( \lambda_1(A) \) and \( \lambda_2(A) \) denote the smallest and largest eigenvalue of the matrix \( A \), respectively. For \( \mathcal{C}_n, \, n \geq 1 \),
\[ \frac{\sqrt{\lambda_2(C_n)}}{\lambda_1(C_n)} = o(1), \quad \text{as } n \to \infty \]  

(3.4.16)

\[ \frac{\lambda_2(C_n)}{\lambda_1(C_n)} = O(1), \quad \text{as } n \to \infty. \]  

(3.4.17)

Then,

\[ |\hat{\theta}_n - \theta| \overset{P}{\to} 0, \quad \text{as } n \to \infty. \]  

(3.4.18)

[Note that (3.4.17) and \( \lim_{n \to \infty} \lambda_1(C_n) = +\infty \) imply (3.4.16)].

Proof. Denote, for \( y \in \mathbb{R}^p, \delta \geq 0, \)

\[ N_n(y, \delta) = \frac{1}{\lambda_1(C_n)} \sum_{i=1}^{n} u^\tau c_i \psi(Y_i - c_i^\tau y\delta), \]  

(3.4.19)

then,

\[ N_n(\varepsilon_n, \|\hat{\theta}_n - \theta\|) = 0, \]

where \( \varepsilon_n = (\hat{\theta}_n - \theta)/\|\hat{\theta}_n - \theta\| \). First, we will show that, for any fixed real number \( 0 \leq \delta < 1, \) as \( n \to \infty \)

\[ \sup_{\|\varepsilon\|=1} |N_n(\varepsilon, \delta) - E(N_n(\varepsilon, \delta))| \overset{P}{\to} 0. \]  

(3.4.20)

For any fixed \( 0 \leq \delta < 1, \) let

\[ R_{nr}(\varepsilon, \delta) = N_n(I_{e\varepsilon}, \delta) - E(N_n(I_{e\varepsilon}, \delta)), \]  

(3.4.21)
where $\varepsilon \in E_p$, $1 \leq r \leq 2^P$ and $I_r$ is given by (3.3.24). Then,

$$
\sup_{\|\varepsilon\|=1} |N_n(\varepsilon, \delta) - E\{N_n(\varepsilon, \delta)\}| \leq \max_{1 \leq r \leq 2^P} \{ \sup_{\varepsilon \in E_p} |R_{nr}(\varepsilon, \delta)| \},
$$

and (3.4.20) follows from showing that, for any $1 \leq r \leq 2^P$, as $n \to \infty$

$$
\sup_{\varepsilon \in E_p} |R_{nr}(\varepsilon, \delta)| \overset{P}{\to} 0. \tag{3.4.22}
$$

Note that $R_{nr}(0, \delta) = 0$, and, for any $\varepsilon \in E_p$, $E\{R_{nr}(\varepsilon, \delta)\} = 0$. Since, for any $u \in E_p$ and $v \in E_p$,

$$
\text{Var}\{R_{nr}(u, \delta) - R_{nr}(v, \delta)\}
$$

$$
= \frac{1}{[\lambda_1(C_n)]^2} \sum_{i=1}^{\mathcal{R}} [(u \tau I_r \varepsilon_i \psi(t - \varepsilon_i^T I_r v - \delta) - v \tau I_r \varepsilon_i \psi(t - \varepsilon_i^T I_r v - \delta)]^2 dF(t)
$$

$$
\leq \frac{M_1}{[\lambda_1(C_n)]^2} \sum_{i=1}^{\mathcal{R}} [(u - v) \tau I_r \varepsilon_i]^2
$$

$$
= \frac{M_1}{[\lambda_1(C_n)]^2} (u - v) \tau I_r C_n I_r (u - v)
$$

$$
\leq \frac{M_1 \lambda_2(C_n)}{[\lambda_1(C_n)]^2} \|u - v\|^2,
$$

where $M_1$ is a constant, then, from the same argument used in the proof of (3.3.19) of Theorem 3.3.2, we have, for arbitrary $\rho > 0$

$$
P\{ \sup_{u \in E_p} |R_{nr}(u, \delta)| \geq \rho \} \leq \frac{M_2}{\rho^2} \left\{ \frac{\lambda_2(C_n)}{[\lambda_1(C_n)]^2} \right\}, \tag{3.4.23}
$$
where $M_2$ is a constant. Therefore, (3.4.22) follows from (3.4.16) and (3.4.23).

Since, for an arbitrary $\rho>0$, by (3.3.22), (3.4.17), (N2), (N4) and DCT, there exists a real number $\delta$ such that $0<\delta<\rho$ and, for all $n\geq 1$,

\[
\sup_{\|\varepsilon\|_1=1} \left| E\{N_n(\varepsilon, \delta)\} + \frac{\delta \gamma_0}{\lambda_1(C_n)} \varepsilon^T C_n \varepsilon \right|
\]

\[
= \sup_{\|\varepsilon\|_1=1} \left| \frac{1}{\lambda_1(C_n)} \left( \sum_{i=1}^n \varepsilon_i \sum_{i=1}^n \psi(y - \varepsilon_i \tau \varepsilon_\delta) dF(y) + \delta \gamma_0 \varepsilon^T C_n \varepsilon \right) \right|
\]

\[
= \sup_{\|\varepsilon\|_1=1} \left| \frac{1}{\lambda_1(C_n)} \left( \sum_{i=1}^n \varepsilon_i \sum_{i=1}^n \left[F(y) - F(y + \varepsilon_i \tau \varepsilon_\delta)\right] d\psi(y) + \delta \gamma_0 \varepsilon^T C_n \varepsilon \right) \right|
\]

\[
= \sup_{\|\varepsilon\|_1=1} \left| \frac{\delta \lambda_2(C_n)}{\lambda_1(C_n)} \sum_{i=1}^n \left(\varepsilon_i \tau \varepsilon_\delta\right)^2 \left[F'(\varepsilon_i) - F'(y)\right] d\psi(y) \right|
\]

\[
\leq \frac{\delta \lambda_2(C_n)}{\lambda_1(C_n)} \int \sup_{\|\varepsilon\|_1=1} \left| F'(\varepsilon_i) - F'(y)\right| d\psi(y)
\]

\[
\leq \delta M_3 \int \sup_{\|\varepsilon\|_1=1} \left| F'(\varepsilon_i) - F'(y)\right| d\psi(y) \leq \frac{\delta \gamma_0}{2}
\]

where $\xi$ is between $y$ and $(y + \varepsilon_i \tau \varepsilon_\delta)$, and $M_3$ is a constant. Therefore, since $N_n(\varepsilon, \delta)$ is nonincreasing in $\delta$, we have

\[
P\{|\hat{\Phi}_n - \theta| > \rho\} \leq P\{|\hat{\Phi}_n - \theta| > \delta\} \leq P\{N_n(\varepsilon_n, \delta) \geq 0\}
\]

\[
= P\{N_n(\varepsilon_n, \delta) + \frac{\delta \gamma_0}{\lambda_1(C_n)} \varepsilon^T n C_n \varepsilon_n \geq \frac{\delta \gamma_0}{\lambda_1(C_n)} \varepsilon^T n C_n \varepsilon_n \}
\]
\[
\leq P \left\{ \frac{\delta \gamma_0}{\lambda_1(C_n)} |e_n^T C_n e_n \geq \delta \gamma_0 \right\}
\]
\[
\leq P \left\{ \sup_{\|e\|=1} |N_n(e) - E[N_n(e)]| + \sup_{\|e\|=1} |E[N_n(e)] + \frac{\delta \gamma_0}{\lambda_1(C_n)} e^T C_n e| \geq \frac{\delta \gamma_0}{2} \right\}.
\] (3.4.24)

Hence, (3.4.18) follows from (3.4.20) and (3.4.24). \(\Box\)

An alternative set of conditions on \(\psi\) and \(F\) is given next for the weak consistency of the M-estimators of regression, where we require weaker conditions on \(F\), but stronger conditions on \(\psi\).

**Theorem 3.4.3.** Assume that, in addition to (N4), (N5), and (C1) through (C3), \(\psi\) and \(F\) satisfy the following conditions:

(N1') \(\psi\) is differentiable with a bounded \(\psi'\)

(N2') \(P\{Y_1 \in \{\text{discontinuity points of } \psi'\}\} = 0\).

Then, as \(n \to \infty\)

\[
|\hat{\beta}_n - \beta| \overset{P}{\to} 0.
\] (3.4.25)

**Proof.** For an arbitrary \(\rho > 0\), by (C2), (3.4.17), (N1'), (N2'), (N4) and DCT, there exists a real number \(\delta\) such that \(0 < \delta < \rho\) and
\[
\sup_{\|\varepsilon\| = 1} \left| E(N_n(\varepsilon, \delta)) + \frac{\delta \gamma}{\lambda_1(C_n)} \varepsilon^\top C_n \varepsilon \right|
\]

\[
= \sup_{\|\varepsilon\| = 1} \left| \frac{1}{\lambda_1(C_n)} \sum_{i=1}^{\mathcal{B}} (\varepsilon^\top \varepsilon_i)^2 \left[ (\psi'(\varepsilon) - \psi'(y)) dF(y) + \delta \gamma \varepsilon^\top C_n \varepsilon \right] \right|
\]

\[
\leq \frac{\delta \lambda_2(C_n)}{\lambda_1(C_n)} \int \sup_{\|\varepsilon\| = 1} |\psi'(\varepsilon) - \psi'(y)| dF(y)
\]

\[
\leq \delta M_3 \int \sup_{\|\varepsilon\| = 1} |\psi'(\varepsilon) - \psi'(y)| dF(y) \leq \frac{\delta \gamma}{2}
\]

where \( \varepsilon \) is between \( y \) and \( (y - c_i^T \hat{\varepsilon}) \), and \( M_3 \) is a constant. Therefore, by the proof of Theorem 3.4.2, (3.4.25) follows from (3.4.16) and

\[
P\left\{ \| \hat{\sigma}_n - \sigma \| > \rho \right\} \leq P\left\{ \sup_{\|\varepsilon\| = 1} |N_n(\varepsilon, \delta) - E\{N_n(\varepsilon, \delta)\}| \geq \frac{\delta \gamma}{2} \right\}
\]

\[
\leq \sum_{i=1}^{\mathcal{B}} P\left\{ \sup_{\varepsilon \in \mathcal{E}} |R_{nr}(\varepsilon, \delta)| \geq \frac{\delta \gamma}{2} \right\} \leq \frac{M_4}{(\delta \gamma)^2} \left\{ \frac{\lambda_2(C_n)}{\lambda_1(C_n)^2} \right\}, \tag{3.4.26}
\]

where \( M_4 \) is a constant.

In the following two theorems, we will study the strong consistency of the M-estimators of regression.

**THEOREM 3.4.4.** Assume that, in addition to the assumptions of
Theorem 3.4.2, \( \{\lambda_1(C_n)\} \) is nondecreasing in \( n \) and that there exists an increasing subsequence \( \{n_k\} \) of \( \{n\} \) such that

\[
\sum_{k=1}^{\infty} \frac{\lambda_2(C_{n_k})}{\left( \frac{\lambda_1(C_{n_k})}{\lambda_1(C_n)} \right)^2} < \infty, \tag{3.4.27}
\]

and

\[
\sum_{k=1}^{n_k+1-n_k} \sum_{q=1}^{\lambda_2(C_{n_k})-\lambda_1(C_{n_k})} \left( \frac{\lambda_2(C_{n_k+q})-\lambda_2(C_{n_k})}{\left( \frac{\lambda_1(C_{n_k+q})}{\lambda_1(C_{n_k})} \right)^2} + \frac{\lambda_2(C_{n_k})-\lambda_1(C_{n_k})}{\left( \frac{\lambda_1(C_{n_k})}{\lambda_1(C_n)} \right)^2} \right) < \infty. \tag{3.4.28}
\]

Then, as \( n \to \infty \)

\[
|\hat{\beta}_n - \beta| \to 0, \quad \text{a.s.} \tag{3.4.29}
\]

Proof. From the proof of Theorem 3.4.2, (3.4.23) and (3.4.24) imply that, for arbitrary \( \rho > 0 \), there exists \( \delta > 0 \) such that

\[
P]\|\hat{\beta}_n - \beta\| > \rho ) \leq P \left\{ \sup_{\|e\|=1} |N_n(e, \delta) - E\{N_n(e, \delta)\}| \geq \frac{\epsilon}{2} \right\}
\leq \sum_{r=1}^{p} P \left\{ \sup_{e \in E_{\rho}} |R_{nr}(e, \delta)| \geq \frac{\epsilon}{2} \right\} \leq \frac{M_3}{(\delta \gamma_0)^2} \left\{ \frac{\lambda_2(C_n)}{\lambda_1(C_n)^2} \right\}, \tag{3.4.30}
\]

where \( M_3 \) is a constant. Hence, by (3.4.27), we have, for arbitrary \( \epsilon > 0 \)

\[
\sum_{k=1}^{\infty} P \left\{ \sup_{\|e\|=1} |N_{n_k}(e, \delta) - E\{N_{n_k}(e, \delta)\}| \geq \epsilon \right\} \leq \frac{4M_3}{\epsilon^2} \sum_{k=1}^{\infty} \frac{\lambda_2(C_{n_k})}{\left( \frac{\lambda_1(C_{n_k})}{\lambda_1(C_n)} \right)^2} < \infty.
\]
Therefore, as \( k \to \infty \)

\[
\sup_{\|e\|=1} \left| N_{n_k}(e, \delta) - \mathbb{E}\{N_{n_k}(e, \delta)\} \right| \to 0, \quad \text{a.s.} \quad (3.4.31)
\]

Let

\[
S_q(e) = \sum_{i=1}^{q} \xi_i = \lambda_1(C_q)\{N_q(e, \delta) - \mathbb{E}\{N_q(e, \delta)\}\}, \quad q \geq 1
\]

(3.4.32)

where \( \xi_i = e^T \xi_i \psi(Y_i - c_i^T e \delta) - e^T \xi_i \mathbb{E}\{\psi(Y_i - c_i^T e \delta)\} \), and let

\[
\eta_q(e) = S_{n_k+q}(e) - S_{n_k}(e), \quad 1 \leq q \leq (n_{k+1} - n_k)
\]

(3.4.33)

\[
D_k(e) = \max_{n_k \leq q \leq n_{k+1}} \left\{ |S_q(e) - S_{n_k}(e)| \right\}
\]

\[
= \max_{1 \leq q \leq (n_{k+1} - n_k)} \{ |\eta_q(e)| \}.
\]

(3.4.34)

Then, \( \mathbb{E}\{\eta_q(e)\} = 0 \), \( \eta_q(0) = 0 \), and, for any \( u, v \in \mathbb{R}^p \)

\[
\text{Var}\{\eta_q(u) - \eta_q(v)\}
\]

\[
= \sum_{i=n_k+1}^{n_k+q} \left\{ \left[ u^T \xi_i \psi(t - c_i^T u \delta) - v^T \xi_i \psi(t - c_i^T v \delta) \right]^2 \right\} \mathbb{d}F(t)
\]

\[
\leq M_1 \sum_{i=n_k+1}^{n_k+q} \{(u - v)^T \xi_i \}^2
\]

\[
= M_1 (u - v)^T (C_{n_k+q} - C_{n_k})(u - v)
\]
\[ \leq M_1 \left\{ \lambda_2(C_{n_k+q}) - \lambda_1(C_{n_k}) \right\} \|u - y\|^2 \]

where \( M_1 \) is a constant. Using the same argument referring to (3.4.23), we have, for any \( \epsilon > 0 \) and \( 1 \leq q \leq (n_{k+1} - n_k) \),

\[ P\left\{ \sup_{\|e\|=1} |\eta_q(e)| \geq \epsilon \right\} \leq \frac{M_4}{\epsilon^2} \left\{ \lambda_2(C_{n_k+q}) - \lambda_1(C_{n_k}) \right\}, \quad (3.4.35) \]

where \( M_4 \) is a constant. Note that for symmetric matrices \( A \) and \( B \) (of the same order), \( \lambda_2(A + B) \leq \lambda_2(A) + \lambda_2(B) \) [viz., Rao (1965, page 55)]. Hence, by (3.4.35) and (3.4.28), we have that, for arbitrary \( \epsilon > 0 \),

\[ \sum_{k=1}^{\infty} P\left\{ \sup_{\|e\|=1} \frac{|D_k(e)|}{\lambda_1(C_{n_k})} \geq \epsilon \right\} \leq \sum_{k=1}^{\infty} \sum_{q=1}^{n_{k+1} - n_k} P\left\{ \sup_{\|e\|=1} |\eta_q(e)| \geq \epsilon \lambda_1(C_{n_k}) \right\} \]

\[ \leq \frac{M_4}{\epsilon^2} \sum_{k=1}^{\infty} \sum_{q=1}^{n_{k+1} - n_k} \left( \frac{\lambda_2(C_{n_k+q}) - \lambda_1(C_{n_k})}{\lambda_1(C_{n_k})^2} \right) \]

\[ \leq \frac{M_4}{\epsilon^2} \sum_{k=1}^{\infty} \sum_{q=1}^{n_{k+1} - n_k} \left( \frac{\lambda_2(C_{n_k+q}) - C_{n_k}}{\lambda_1(C_{n_k})^2} + \frac{\lambda_2(C_{n_k}) - \lambda_1(C_{n_k})}{\lambda_1(C_{n_k})^2} \right) < \infty. \]

Therefore, as \( k \to \infty \)

\[ \sup_{\|e\|=1} \frac{|D_k(e)|}{\lambda_1(C_{n_k})} \to 0, \quad \text{a.s.} \quad (3.4.36) \]

Since, for \( n_k \leq q \leq n_{k+1} \),
\[
\frac{|S_q(\epsilon)|}{\lambda_1(C_q)} \leq \frac{|S_{n_k}(\epsilon)| + |D_k(\epsilon)|}{\lambda_1(C_{n_k})} = \left| N_{n_k}(\epsilon, \delta) - E\{N_{n_k}(\epsilon, \delta)\}\right| + \frac{|D_k(\epsilon)|}{\lambda_1(C_{n_k})}
\]

therefore, by (3.4.31) and (3.4.36), we have, as \( n \to \infty \)

\[
\sup_{\|\epsilon\| = 1} \frac{|S_n(\epsilon)|}{\lambda_1(C_n)} = \sup_{\|\epsilon\| = 1} \left| N_n(\epsilon, \delta) - E\{N_n(\epsilon, \delta)\}\right| \to 0, \quad \text{a.s.} \quad (3.4.37)
\]

Therefore, by (3.4.30) and (3.4.37), (3.4.29) follows from

\[
P\left\{ \lim_{n \to \infty} \{ \|\hat{\beta}_n - \beta\| > \rho \} \right\} = P\left\{ \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \{ \|\hat{\beta}_n - \beta\| > \rho \} \right\}
\leq P\left\{ \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \left\{ \sup_{\|\epsilon\| = 1} |N_n(\epsilon, \delta) - E\{N_n(\epsilon, \delta)\}| \geq \frac{\delta \gamma_0}{2} \right\} \right\}
\]

\[
= P\left\{ \lim_{n \to \infty} \left\{ \sup_{\|\epsilon\| = 1} |N_n(\epsilon, \delta) - E\{N_n(\epsilon, \delta)\}| \geq \frac{\delta \gamma_0}{2} \right\} \right\} = 0.
\]

Next theorem gives an alternative set of conditions on \( \psi \) and \( F \) for strong consistency of the M-estimators of regression, where, corresponding to Theorem 3.4.3, we require weaker conditions on \( F \), but stronger conditions on \( \psi \).

**THEOREM 3.4.5.** Assume that, in addition to the assumptions of Theorem 3.4.3, \( \{\lambda_1(C_n)\} \) is nondecreasing in \( n \) and that (3.4.27) and (3.4.28) hold for an increasing subsequence \( \{n_k\} \) of \( \{n\} \). Then, as \( n \to \infty \)

\[
|\hat{\beta}_n - \beta| \to 0, \quad \text{a.s.} \quad (3.4.38)
\]
Proof. Using the proof of Theorem 3.4.3, the proof is similar to the one of Theorem 3.4.4. \(\square\)

As we can see, we require stronger conditions on \(\psi\) and \(F\) for the strong consistency so that we have (3.4.23), which implies the weak consistency. Although it may not provide the bound of (3.4.23), the Hadamard differentiability approach allows weaker conditions for the weak consistency of the M-estimator of regression in the literature, because the bound of (3.4.23) is not necessary for weak consistency concern.
CHAPTER IV

ON THE SECOND-ORDER HADAMARD DIFFERENTIABILITY
AND ITS APPLICATION TO THE UNIFORM ASYMPTOTIC LINEARITY
OF M-ESTIMATION IN LINEAR MODELS

4.1 Introduction

Consider the general linear model:

\[ X_i = \beta^\tau c_i + e_i, \quad i \geq 1 \]  \hspace{1cm} (4.1.1)

where the \( c_i \) are known \( p \)-vectors of regression constants, \( \beta = (\beta_1, \ldots, \beta_p)^\tau \) is the vector of unknown (regression) parameters, \( p \geq 1 \), and is to be estimated from \( n \) observations \( X_1, \ldots, X_n \), and \( e_i \) are independent and identically distributed random variables (i.i.d.r.v.) with distribution function (d.f.) \( F \). Based on a suitable score function \( \psi: \mathbb{R} \rightarrow \mathbb{R} \), an (robust) M-estimator \( \hat{\beta}_n \) of \( \beta \) is defined as a solution (with respect to \( \theta \)) of the following equations:

\[ \sum_{i=1}^{n} c_i \psi(X_i - c_i \tau \theta) = 0. \]  \hspace{1cm} (4.1.2)

Setting \( Y_i = X_i - \beta^\tau c_i \) (i.i.d.r.v's with d.f. \( F \)), and
\[ C_n = \sum_{i=1}^{p} c_i c_i^\top = (r_{ni})_{1 \leq i, j \leq p} \quad (4.1.3) \]

\[ C_n^o = \text{Diag}(\sqrt{r_{n11}}, \ldots, \sqrt{r_{npn}}) \quad (4.1.4) \]

\[ \varepsilon_{ni} = (C_n^o)^{-1} \varepsilon_i = (c_{ni1}, \ldots, c_{nip})^\top, \quad 1 \leq i \leq n \quad (4.1.5) \]

\[ Q_n = \sum_{i=1}^{n} \varepsilon_{ni} \varepsilon_{ni}^\top = (C_n^o)^{-1} C_n (C_n^o)^{-1} \quad (4.1.6) \]

\[ u = C_n^o (\theta - \varrho) \in \mathbb{R}^p \quad (4.1.7) \]

then, for any \( 1 \leq j \leq p \),

\[ \| \varepsilon_{n,j} \|^2 = 1 \quad (4.1.8) \]

where \( \| \cdot \| \) stands for the Euclidean norm and

\[ \varepsilon_{n,j} = (c_{n1j}, \ldots, c_{nnj})^\top, \quad (4.1.9) \]

and (4.1.2) is equivalent to

\[ M_n(u) = \emptyset \quad \text{(with respect to } u) \quad (4.1.10) \]

where

\[ M_n(u) = \sum_{i=1}^{n} \varepsilon_{ni} \psi(Y_i - \varepsilon_{ni}^\top u), \quad (4.1.11) \]
i.e.,

\[ M_n(\hat{u}_n) = 0 \]  \hspace{1cm} (4.1.12)

for \( \hat{u}_n = C_n^0 (\hat{\beta}_n - \beta) \). Under certain regularity conditions on \( \{c_i\} \), \( \psi \) and \( F \), we, by incorporating the first-order Hadamard differentiability, have shown (in Chapter III) the uniform asymptotic linearity of the M-estimators of regression, i.e., for an arbitrary \( K > 0 \), as \( n \to \infty \)

\[
\sup_{|u| \leq K} \left| M_n(u) - M_n(0) + Q_n u \gamma \right| \xrightarrow{P} 0, \hspace{1cm} (4.1.13)
\]

where \( |x| = |(x_1, \ldots, x_q)| = \max_{1 \leq i \leq q} |x_i| \), and \( \gamma = \int \psi'(x) dF(x) \). The results of Chapter III shows that (4.1.13) provides an easy access to the study of the asymptotic properties of the M-estimators of regression \( \hat{\beta}_n \).

Since (4.1.13) is shown by incorporating the first-order Hadamard differentiability, we naturally consider using the second-order Hadamard differentiability, which provides a second order expansion, i.e., a more precise approximation, to study the more detailed asymptotic property of (4.1.13): the convergence rate in probability of (4.1.13), which will lead to an asymptotic representation of the M-estimators of regression.

Jurečková and Sen (1984) studied the convergence rate in probability of (4.1.13) and gave an asymptotic representation of M-estimators in linear models, assuming that \( F \) has finite Fisher’s information and that \( \psi \) is a nondecreasing and
continuous function with a bounded derivative and \( \psi' = 0 \) outside of a finite interval.

The interest of our current study is to study the convergence rate in probability of (4.1.13) by incorporating the second-order Hadamard differentiability and give an asymptotic representation of the M-estimator of regression under the weak conditions on \( \psi \) and \( F \).

Some notations along with basic assumptions are presented in Section 4.2. The general definition of the second-order Hadamard differentiability and some related theoretical results are given in Section 4.3. The results on the convergence rate in probability of (4.1.13) by incorporating the second-order Hadamard differentiability and an asymptotic representation of the M-estimator of regression are derived in Section 4.4, where we basically require weaker conditions on \( \{c_i\} \) for \( 1 \leq i \leq n \), \( \psi \) and \( F \) than Jurečková and Sen (1984).
4.2 Notations and Assumptions

Since $M_n(u)$ involves the following empirical function

$$S_n^*(t,u)=\sum_{i=1}^n c_{ni} I(Y_i \leq F^{-1}(t)+\epsilon_{ni}^T u), \quad t \in [0,1], \ u \in \mathbb{R}^p$$  \hspace{1cm} (4.2.1)

particularly, for each $u \in \mathbb{R}^p$ and $1 \leq k \leq p$, $M_{nk}(u)$ (the $k$-th component of $M_n(u)$) is a linear functional of

$$S_{nk}^*(t,u)=\sum_{i=1}^n c_{nik}^* I(Y_i \leq F^{-1}(t)+\epsilon_{ni}^T u)$$  \hspace{1cm} (4.2.2)

(the $k$-th component of $S_n^*(t,u)$), viz.,

$$M_{nk}(t,u) = \int \psi(F^{-1}(t)) \, dS_{nk}^*(t),$$ \hspace{1cm} (4.2.3)

we consider the expected value of $S_{nk}^*(t,u)$:

$$S_{nk}(t,u)=\sum_{i=1}^n c_{nik} F(F^{-1}(t)+\epsilon_{ni}^T u).$$  \hspace{1cm} (4.2.4)

We also write for every $1 \leq i \leq n, \ 1 \leq k \leq p, \ t \in [0,1]$ and $u \in \mathbb{R}^p$,

$$c_{nik}=c_{nik}^+-c_{nik}^-, \quad c_{nk}^+=\max\{0, c_{nik}\}, \quad c_{nik}^-=\min\{0, c_{nik}\};$$  \hspace{1cm} (4.2.5)

$$\epsilon_{ni}=\epsilon_{ni}^+-\epsilon_{ni}^-, \quad \epsilon_{ni}^+=(c_{ni1}^+, \ldots, c_{nip}^+)^T, \quad \epsilon_{ni}^-=(-c_{ni1}^-, \ldots, -c_{nip}^-)^T;$$  \hspace{1cm} (4.2.6)
\[ \zeta_{n,k} = \zeta_{n,k}^+ - \zeta_{n,k}^- \]

\[ \zeta_{n,k}^+ = (c_{n1k}^+, \ldots, c_{n nk}^+)^\tau, \quad \zeta_{n,k}^- = (c_{n1k}^-, \ldots, c_{n nk}^-)^\tau; \]

\[ b_{nk}^+ = \sum_{i=1}^{n} c_{nik}^+, \quad b_{nk}^- = \sum_{i=1}^{n} c_{nik}^-; \]

\[ b_n^+ = (b_{n1}^+, \ldots, b_{np}^+)^\tau, \quad b_n^- = (b_{n1}^-, \ldots, b_{np}^-)^\tau; \]

\[ (d_{nk}^+)^2 = ||c_{nk}^+||^2, \quad (d_{nk}^-)^2 = ||c_{nk}^-||^2; \]

\[ \bar{c}_{nik}^+ = \begin{cases} \frac{c_{nik}^+}{d_{nk}^+} & \text{if } d_{nk}^+ > 0 \\ 0 & \text{if } d_{nk}^+ = 0 \end{cases}, \quad \bar{c}_{nik}^- = \begin{cases} \frac{c_{nik}^-}{d_{nk}^-} & \text{if } d_{nk}^- > 0 \\ 0 & \text{if } d_{nk}^- = 0; \end{cases} \]

\[ \bar{\zeta}_{ni}^+ = (\bar{\zeta}_{ni1}^+, \ldots, \bar{\zeta}_{nip})^\tau, \quad \bar{\zeta}_{ni}^- = (\bar{\zeta}_{ni1}^-, \ldots, \bar{\zeta}_{nip})^\tau; \]

\[ S_{nk}^{**}(t,y) = \sum_{i=1}^{n} c_{nik}^+ I(Y_i \leq F^{-1}(t) + \zeta_{ni}^+ y), \quad S_{nk}^{**}(t,y) = E\{S_{nk}^{**}(t,y)\}; \]

\[ S_{nk}^{**}(t,y) = \sum_{i=1}^{n} c_{nik}^- I(Y_i \leq F^{-1}(t) + \zeta_{ni}^- y), \quad S_{nk}^{**}(t,y) = E\{S_{nk}^{**}(t,y)\}; \]

so that

\[ \sum_{i=1}^{n} c_{nik} = b_{nk}^+ - b_{nk}^-; \]

\[ S_{nk}^{**}(t,y) = S_{nk}^{**}(t,y) - S_{nk}^{**}(t,y), \quad S_{nk}(t,y) = S_{nk}^{+}(t,y) - S_{nk}^{-}(t,y); \]

and
\[ \| \varepsilon_{n,k} \|^2 = (d_{nk}^+)^2 + (d_{nk}^-)^2 = 1. \] (4.2.17)

In this chapter, we will always consider the D[0,1] space (of right continuous real valued functions with left hand limits) is endowed with the uniform topology. It is easy to see that, for every \( y \in \mathbb{R}^p \) and \( 1 \leq k \leq p \), \( S_{nk}^*(\cdot, y) \) is an element of D[0,1]. A convention which we will follow through the whole chapter is: a function \( f: \mathbb{R} \rightarrow \mathbb{R} \) is nondecreasing if \( f(x) \leq f(y) \) for \( x \leq y \), and is increasing if \( f(x) < f(y) \) for \( x < y \). Analogously, for nonincreasing and decreasing.

Some assumptions, which may be required for our results in this chapter, are given as below:

(A1) \( c_{n,k} \geq 0 \), \( i=1, 2, \ldots, n \)

(A2) \[ \lim_{n \to \infty} n \max_{1 \leq i \leq n} \| c_{ni} \|^2 < \infty \]

(A3) \[ \lim_{n \to \infty} n \max_{1 \leq i \leq n} \| \varepsilon_{ni}^+ \|^2 < \infty, \quad \lim_{n \to \infty} n \max_{1 \leq i \leq n} \| \varepsilon_{ni}^- \|^2 < \infty \]

(A4) \[ \inf_{n \geq 1} \lambda_1(Q_n) > 0 \], where \( \lambda_1(Q_n) \) denotes the smallest eigenvalue of the matrix \( Q_n \)

(B1) \( F \) is absolutely continuous with a derivative \( F' \) which is positive and continuous with limits at \( \pm \infty \)

(B2) \( F'' \) exists and is bounded in any finite interval
(C1) $0 < \gamma = \int \psi' \, dF < \infty$

(C2) $\psi$ is bounded and absolutely continuous with $|\psi| \leq M$, a.e.,
where $M$ is a constant, and $\psi' = 0$ outside of a finite interval

(C3) $\psi$ is a nondecreasing function with a range including positive and
negative real numbers

(C4) $\int \psi \, dF = 0$.

We notice that, for any fixed $k=1, \ldots, p$, (A1) and (A2) imply that

$$\sum_{i=1}^{n} c_{nik} \rightarrow \infty, \quad \text{as } n \rightarrow \infty \tag{4.2.18}$$

and (A2) implies two facts: there exists $C > 0$ such that for $n \geq 1$ and any $1 \leq j \leq p$,

$$\left( \sum_{i=1}^{n} |c_{nij}| \right)^{a_{n}} \leq C, \tag{4.2.19}$$

where

$$a_{n} = \max_{1 \leq j \leq p} \max_{1 \leq i \leq n} |c_{nij}|, \tag{4.2.20}$$

and

$$\max_{1 \leq i \leq n} c_{nij}^{2} \rightarrow 0, \quad \text{as } n \rightarrow \infty. \tag{4.2.21}$$
We also notice that (B1) implies that $F'$ is bounded and uniformly continuous.

4.3 On the Second-Order Hadamard Differentiability

The first-order Hadamard differentiability is usually defined as the following: Let $V$ and $W$ be the topological vector spaces and $L_1(V, W)$ be the set of continuous linear transformations from $V$ to $W$. Let $\mathcal{A}$ be an open set of $V$.

**DEFINITION.** A functional $T: \mathcal{A} \rightarrow W$ is *Hadamard Differentiable* (or *Compact Differentiable*) at $F \in \mathcal{A}$ if there exists $T'_F \in L_1(V, W)$ such that for any compact set $\Gamma$ of $V$,

$$
\lim_{t \to 0} \frac{T(F+tH) - T(F) - T'_F(tH)}{t} = 0
$$

(4.3.1)

uniformly for any $H \in \Gamma$. The linear function $T'_F$ is called the *Hadamard Derivative* of $T$ at $F$.

For convenience sake, in (4.3.1), we usually denote

$$
\text{Rem}_1(tH) = T(F+tH) - T(F) - T'_F(tH)
$$

(4.3.2)

as the remainder term of the first order expansion.
Naturally, let \( \mathcal{C}(V, W) \) be the set of continuous functionals from \( V \) to \( W \), and let

\[
L_2(V, W) = \{ f; f \in \mathcal{C}(V, W), f(tH) = t^2 f(H) \text{ for any } H \in V, t \in \mathbb{R} \},
\]

we define the second-order Hadamard differentiability as below.

**DEFINITION.** A functional \( T: \mathcal{A} \rightarrow W \) is Second-Order Hadamard Differentiable at \( F \in \mathcal{A} \) if there exist \( T'_F \in L_1(V, W) \) and \( T''_F \in L_2(V, W) \) such that for any compact set \( \Gamma \) of \( V \),

\[
\lim_{t \rightarrow 0} \frac{T(F+tH)-T(F)-T'_F(tH)-\frac{1}{2}T''_F(tH)}{t^2} = 0
\]

uniformly for any \( H \in \Gamma \). \( T'_F \) and \( T''_F \) are called the First- and Second-Order Hadamard Derivative of \( T \) at \( F \), respectively.

We also denote the remainder term of the second order expansion as the following:

\[
\text{Rem}_2(tH) = T(F+tH) - T(H) - T'_F(tH) - T''_F(tH).
\]

**Remark (1).** Our definition of the second-order Hadamard differentiability is consistent with the one given by Sen (1988). Also, we have
\[ T_F'(\delta_x - F) = \text{IC}(x; F, T) = \frac{d}{dt} T(F + t(\delta_x - F)) \bigg|_{t=0} \quad (4.3.6) \]

and

\[ T_F''(\delta_x - F) = \frac{d^2}{dt^2} T(F + t(\delta_x - F)) \bigg|_{t=0} \quad (4.3.7) \]

where \( \delta_x \) is the d.f. of the point mass one at \( x \).

Remark (2). From our definition of the second-order Hadamard differentiability, it is obvious that the existence of the second-order Hadamard derivative implies the existence of the first-order Hadamard derivative.

It is known that the chain rule holds for the first-order Hadamard differentiability (Fernholz, 1983), which makes it useful. We will show that the chain rule also holds for the second-order Hadamard differentiability.

**PROPOSITION 4.3.1.** Let \( V, W \) and \( Z \) be the topological vector spaces with \( T: V \rightarrow W \) and \( Q: W \rightarrow Z \). If \( T \) is second-order Hadamard differentiable at \( F \epsilon V \) and if \( Q \) is second-order Hadamard differentiable at \( T(F) \epsilon W \), then \( \tau = Q \circ T \) is second-order Hadamard differentiable at \( F \) and

\[ \tau'_F = Q'_T(F) \circ T'_F \quad (4.3.8) \]

\[ \tau''_F = (Q \circ T)'_F = Q''_T(F) \circ T'_F + Q'_T(F) \circ T''_F. \quad (4.3.9) \]
Proof. Since $T$ and $Q$ are second-order Hadamard differentiable at $F$ and $T(F)$, respectively, for any compact set $\Gamma_V$ of $V$ and compact set $\Gamma_W$ of $W$, we have

$$
\lim_{t \to 0} \frac{T(F+tH)-T(F)-T'_F(tH)-\frac{1}{2}T''_F(tH)}{t^2} = 0 \quad (4.3.10)
$$

uniformly for any $H \in \Gamma_V$, and

$$
\lim_{t \to 0} \frac{Q(T(F)+tG)-Q(T(F))-Q'_T(F)(tG)-\frac{1}{2}Q''_T(F)(tG)}{t^2} = 0 \quad (4.3.11)
$$

uniformly for any $G \in \Gamma_W$.

By Proposition 3.1.2 of Fernholz (1983), we know

$$
\tau'_F = Q'_T(F) \circ T'_F,
$$

and obviously $\tau'_F \in L_1(V, Z)$. It is also obvious that $\tau''_F$ given by (4.3.9) is an element of $L_2(V, Z)$. From (4.3.10) we have

$$
T(F+tH) = T(F) + T'_F(tH) + \frac{1}{2}T''_F(tH) + o(1)t^2,
$$

so, by the linearity of $Q'_T(F)$,

$$
\tau(F+tH) - \tau(F) - \tau'_F(tH)
$$

$$
= Q(T(F+tH)) - Q(T(F)) - Q'_T(F)(T'_F(tH))
$$
\[ Q(T(F)) + T_F'(tH) + \frac{1}{2} T_F''(tH) + o(1)t^2 - Q(T(F)) - Q'_T(F)(T_F'(tH)) \]

\[ = \left\{ Q(T(F)) + t\{ T_F'(H) + \frac{1}{2} T_F''(H) + o(1)t\} - Q(T(F)) - \right. \]

\[ - Q'_T(F)(t\{ T_F'(H) + \frac{1}{2} T_F''(H) + o(1)t\}) \left\} + \right. \]

\[ + \frac{1}{2} t^2 Q''_T(F)(T_F''(H)) + t^2 Q'_T(F)(o(1)). \quad (4.3.12) \]

By (4.3.11) and the uniform continuity of \( Q''_T(F) \) on a compact set, we have, for any compact set \( \Gamma \) of \( V \),

\[ \lim_{t \to 0} \frac{1}{t^2} \left\{ Q(T(F)) + t\{ T_F'(H) + \frac{1}{2} T_F''(H) + o(1)t\} - Q(T(F)) - \right. \]

\[ - Q'_T(F)(t\{ T_F'(H) + \frac{1}{2} T_F''(H) + o(1)t\}) \left\} - \frac{1}{2} Q''_T(F)(T_F'(tH)) \right. \]

\[ = \lim_{t \to 0} \frac{Q''_T(F)(T_F'(H) + \frac{1}{2} T_F''(H) + o(1)t) - Q''_T(F)(T_F'(tH))}{2t^2} \]

\[ = \lim_{t \to 0} \frac{1}{2} \left\{ Q''_T(F)(T_F'(H) + \frac{1}{2} T_F''(H) + o(1)t) - Q''_T(F)(T_F'(H)) \right\} = 0 \quad (4.3.13) \]

uniformly for any \( H \in \Gamma \). Since \( Q'_T(F) \) is continuous, then

\[ \lim_{t \to 0} \frac{t^2 Q'_T(F)(o(1))}{t^2} = \lim_{t \to 0} Q'_T(F)(o(1)) = 0 \quad (4.3.14) \]
uniformly for any \( H \in \Gamma \). Therefore, (4.3.12) through (4.3.14) imply that

\[
\lim_{t \to 0} \frac{\tau(F+tH) - \tau(F) - \tau'_F(tH) - \frac{1}{2} \{Q''_{\tau(F)}(T_{\tau(F)}(tH)) + Q''_{\tau(F)}(T'_{\tau(F)}(tH))\}}{t^2} = 0
\]

(4.3.15)

uniformly for any \( H \in \Gamma \).

\[\square\]

**Proposition 4.3.2.** Let \( L: \mathbb{R} \to \mathbb{R} \) be differentiable and \( L' \) be continuous, bounded and piecewise differentiable with a bounded derivative. Let \( \gamma: \mathbb{D}[0,1] \to L^p[0,1], \ p \geq 1, \) be defined by

\[\gamma(S) = L \circ S,\]

and let \( A \) be the set of points in \( \mathbb{R} \) where \( L' \) is not differentiable. If \( \gamma \) is defined in a neighborhood of \( Q \in \mathbb{D}[0,1] \) and if \( \mu\{x; \ Q(x) \in A\} = 0, \) where \( \mu \) is Lebesgue measure, then \( \gamma \) is second-order Hadamard differentiable at \( Q \) with derivative

\[\gamma'_Q(H) = (L' \circ Q)H\]

and

\[\gamma''_Q(H) = (L'' \circ Q)H^2.\]

**Proof.** For any compact set \( \Gamma \) of \( \mathbb{D}[0,1] \), we need to show that
$$\left\| \frac{\text{Rem}_2(tH)}{t^2} \right\|_{L^p} \to 0$$  \hspace{1cm} (4.3.16)

uniformly for $H \in \Gamma$, as $t \to 0$, where

$$\text{Rem}_2(tH) = L\circ (Q+th) - L\circ Q - (L'\circ Q)tH - \frac{1}{2}(L''\circ Q)t^2H^2.$$

Since $\Gamma$ is compact, for any $\epsilon > 0$, we can choose $H_1, \ldots, H_n \in \Gamma$ such that

$$\inf_{1 \leq i \leq n} \| H - H_i \| < \epsilon$$  \hspace{1cm} (4.3.17)

for any $H \in \Gamma$.

Since for a given $H_i$

$$\frac{\text{Rem}_2(tH_i)(x)}{t^2} =$$

$$= \frac{L(Q(x)+tH_i(x)) - L(Q(x)) - L'(Q(x))tH_i(x) - \frac{1}{2}L''(Q(x))t^2H_i^2(x)}{t^2}$$

$$= \frac{L'(\xi) - L'(Q(x))}{t}H_i(x) - \frac{1}{2}L''(Q(x))H_i^2(x)$$

where $\xi$ is between $Q(x)$ and $(Q(x)+tH_i(x))$, by Lemma 5.4.3 of Fernholz (1983), we have

$$\left| \frac{L'(\xi) - L'(Q(x))}{t} \right| \leq M |H_i(x)|.$$
where $M$ is a bound for $L''$. Therefore, for each $i$,

$$\left| \frac{\mathrm{Rem}_2(tH_i(x))}{t^2} \right| \leq M \ |H(x)|^2$$

where $M$ is a constant. Moreover, for $x$ such that $Q(x) \notin A$,

$$\frac{\mathrm{Rem}_2(tH_i(x))}{t^2} \to 0, \quad \text{as } t \to 0.$$ 

So, by Dominated Convergence Theorem, as $t \to 0$

$$\left\| \frac{\mathrm{Rem}_2(tH_i)}{t^2} \right\|_{L^p} \to 0. \quad (4.3.18)$$

For any $H \in \Gamma$,

$$\left\| \frac{\mathrm{Rem}_2(tH)}{t^2} \right\|_{L^p} \leq \left\| \frac{\mathrm{Rem}_2(tH_i)}{t^2} \right\|_{L^p} + \left\| \frac{\mathrm{Rem}_2(tH)}{t^2} - \frac{\mathrm{Rem}_2(tH_i)}{t^2} \right\|_{L^p} \quad (4.3.19)$$

and

$$\frac{\mathrm{Rem}_2(tH) - \mathrm{Rem}_2(tH_i)}{t^2}$$

$$= \frac{L(Q(x) + tH(x)) - L(Q(x) + tH(x))}{t^2} - \frac{L'(Q(x))(H(x) - H_i(x))}{t}$$

$$- \frac{1}{2}L''(Q(x))(H^2(x) - H_i^2(x))$$
\[ L'(\eta) - L'(Q(x)) = \frac{1}{t} L''(Q(x))\{H(x) - H_1(x)\} - \frac{1}{2} L''(Q(x))\{H^2(x) - H_1^2(x)\}, \]

where \( \eta \) is between \( \{Q(x) + tH(x)\} \) and \( \{Q(x) + tH_1(x)\} \). By Lemma 5.4.3 of Fernholz (1983), we have

\[ \left| \frac{L'(\eta) - L'(Q(x))}{t} \right| \leq M |H(x) - H_1(x)|. \]

So, by (4.3.17),

\[ \inf_{1 \leq i \leq n} \left\| \frac{\text{Rem}_2(tH)}{t^2} - \frac{\text{Rem}_2(tH_1)}{t^2} \right\|_{L^p} \leq M_\Gamma \epsilon \quad (4.3.20) \]

where \( M_\Gamma \) is a constant which depends on \( \Gamma \). Therefore, (4.3.16) follows from (4.3.18) through (4.3.20).

\[ \square \]

**THEOREM 4.3.3.** Suppose \( \tau: D[0,1] \to \mathbb{R} \) is a functional and is second-order Hadamard differentiable at \( U \). For any fixed \( k = 1, 2, \ldots, p \), assume (A1), (A2) and (B1). Then, for any \( K > 0 \), as \( n \to \infty \)

\[ \sup_{|y| \leq K} \left| \left( \sum_{i=1}^{n} c_{n,k} \right) \text{Rem}_2 \left( \frac{S_{n,k}(\cdot, y)}{\sum_{i=1}^{n} c_{n,k}} - U(\cdot) \right) \right|_{L^p} \to 0. \quad (4.3.21) \]

Therefore, we have, as \( n \to \infty \)
\[
\sup_{|y| \leq K} \left| \left( \sum_{i=1}^{n} c_{nik} \right)^2 \left\{ \tau \left( \frac{S_{nk}^*(\cdot, y)}{\sum_{i=1}^{n} c_{nik}^+} \right) - \tau \left( U(\cdot) \right) \right\} - \sum_{i=1}^{n} c_{nik} \tau u \left( S_{nk}^*(\cdot, y) - U(\cdot) \frac{\sum_{i=1}^{n} c_{nik}^+}{c_{nik}} \right) \right. \\
\left. - \frac{1}{2} \tau'' u \left( S_{nk}^*(\cdot, y) - U(\cdot) \frac{\sum_{i=1}^{n} c_{nik}^+}{c_{nik}} \right) \right| \overset{P}{\rightarrow} 0. \tag{4.3.22}
\]

Proof. The proof follows from simply replacing

\[
\frac{\text{Rem}(tH)}{t} \quad \text{by} \quad \frac{\text{Rem}_2(tH)}{t^2}
\]

in the proof of Theorem 2.3.1. □

In the general case of \( \{c_{nik}\} \), we have the following theorem.

**THEOREM 4.3.4.** Suppose \( \tau : D[0,1] \rightarrow \mathbb{R} \) is a functional and is second-order Hadamard differentiable at \( U \). Assume (A3) and (B1). Then, for arbitrary \( 1 \leq k \leq p \) and \( K > 0 \), as \( n \rightarrow \infty \)

\[
\sup_{|y| \leq K} \left| \left( \sum_{i=1}^{n} c_{nik}^+ \right)^2 \left\{ \tau \left( \frac{S_{nk}^*(\cdot, y)}{\sum_{i=1}^{n} c_{nik}^+} \right) - \tau \left( \frac{S_{nk}^*(\cdot, \partial)}{\sum_{i=1}^{n} c_{nik}^+} \right) \right\} - \sum_{i=1}^{n} c_{nik}^+ \tau u \left( S_{nk}^*(\cdot, y) - S_{nk}^*(\cdot, \partial) \right) \right. \\
\left. - \frac{1}{2} \left\{ \tau'' u \left( S_{nk}^*(\cdot, y) - U(\cdot) \frac{\sum_{i=1}^{n} c_{nik}^+}{c_{nik}} \right) \right\} \right| \overset{P}{\rightarrow} 0, \tag{4.3.23}
\]

and
\[
\sup_{|u| \leq K} \left\{ \left( \sum_{i=1}^{n} c_{nik}^* \right)^2 \left\{ \tau \left( \frac{S_{nk}^* (\cdot, u)}{\sum_{i=1}^{n} c_{nik}^*} \right) - \tau \left( \frac{S_{nk}^* (\cdot, 0)}{\sum_{i=1}^{n} c_{nik}^*} \right) - \sum_{i=1}^{n} c_{nik}^* r_{n} \left( S_{nk}^* (\cdot, u) - S_{nk}^* (\cdot, 0) \right) \right\} - \frac{1}{2} \right\} \right\} \mathbb{P} \leq 0.
\]

(4.3.24)

Proof. Using Theorem 4.3.3, by the similar proof of Theorem 2.3.2, we have, for any \(1 \leq k \leq p\) and \(K > 0\), as \(n \to \infty\)

\[
\sup_{|u| \leq K} \left\{ \left( \sum_{i=1}^{n} c_{nik}^* \right)^2 \left\{ \tau \left( \frac{S_{nk}^* (\cdot, u)}{\sum_{i=1}^{n} c_{nik}^*} \right) - \tau \left( \frac{U(\cdot)}{n} \right) \right\} - \sum_{i=1}^{n} c_{nik}^* \tau \left( S_{nk}^* (\cdot, u) - U(\cdot) \sum_{i=1}^{n} c_{nik}^* \right) \right\} - \frac{1}{2} \tau \left( S_{nk}^* (\cdot, u) - U(\cdot) \sum_{i=1}^{n} c_{nik}^* \right) \mathbb{P} \leq 0,
\]

(4.3.25)

and

\[
\sup_{|u| \leq K} \left\{ \left( \sum_{i=1}^{n} c_{nik}^* \right)^2 \left\{ \tau \left( \frac{S_{nk}^* (\cdot, u)}{\sum_{i=1}^{n} c_{nik}^*} \right) - \tau (U(\cdot)) - \sum_{i=1}^{n} c_{nik}^* r_{n} \left( S_{nk}^* (\cdot, u) - U(\cdot) \sum_{i=1}^{n} c_{nik}^* \right) \right\} - \frac{1}{2} \tau \left( S_{nk}^* (\cdot, u) - U(\cdot) \sum_{i=1}^{n} c_{nik}^* \right) \mathbb{P} \leq 0.
\]

(4.3.26)

Therefore, (4.3.23) and (4.3.24) follow from (4.3.25) and (4.3.26), respectively.
4.4 Convergence Rate in Probability of the Uniform Asymptotic Linearity of the M-estimators of Regression

In this section, we define a function \( g: \mathbb{R} \rightarrow \mathbb{R} \) given by

\[
g(x) = \begin{cases} 
F^{-1}(x) & a \leq x \leq b \\
[x-a+F^{-1}(a)F'(F^{-1}(a))]/F'(F^{-1}(a)) & -\infty < x \leq a \\
[x-b+F^{-1}(a)F'(F^{-1}(b))]/F'(F^{-1}(b)) & b \leq x < \infty 
\end{cases}
\] (4.4.1)

where \( 0 < a < b < 1 \), and define a function \( m: \mathbb{R} \rightarrow \mathbb{R} \) given by

\[
m(t) = \psi'(F^{-1}(t)), \quad t \in [0,1].
\] (4.4.2)

So, we have

\[
g'(x) = \begin{cases} 
1/F'(F^{-1}(a)) & -\infty < x \leq a \\
1/F'(F^{-1}(x)) & a \leq x \leq b \\
1/F'(F^{-1}(b)) & b \leq x < \infty,
\end{cases}
\] (4.4.3)

and

\[
\int_{0}^{1} m(t)dt = \int \psi'(x) dF(x) = \gamma.
\] (4.4.4)

Obviously, \( g \) is continuous and differentiable everywhere, and \( g' \) is bounded, continuous and piecewise differentiable with a bounded derivative, if we assume
(B1) and (B2).

**Lemma 4.4.1.** For a fixed \(k=1, \ldots, p\), assume (A1), (A2), (B1), (B2), (C1) and (C2). Then, for any \(K>0\), as \(n \to \infty\)

\[
\sup_{|y| \leq K} \left\{ \left( \sum_{j=1}^{n} c_{njk} \right)^2 \int_0^1 \left\{ g\left( \frac{S_{nk}^*(t,y)}{\sum_{i=1}^{n} c_{nik}} \right) - g\left( \frac{S_{nk}(t,0)}{\sum_{i=1}^{n} c_{nik}} \right) \right\} m(t) dt \right. \\
\left. - \left( \sum_{j=1}^{n} c_{njk} \right) \sum_{i=1}^{n} c_{nik} \epsilon_{ni} \mu \gamma \right\} = O_P(1). \quad (4.4.5)
\]

**Proof.** From the assumptions on \(\psi\) and \(F\), we have that there exist \(a_0, b_0 \in (0,1)\) such that \(m(t)=0\) outside of \([a_0, b_0]\). Choose 'a' and 'b' for (4.4.1) such that \(0 < a < a_0 < b_0 < b < 1\), then, from the properties of \(g\), we have

\[
\left( \sum_{j=1}^{n} c_{njk} \right)^2 \int_0^1 \left\{ g\left( \frac{S_{nk}^*(t,y)}{\sum_{i=1}^{n} c_{nik}} \right) - g\left( \frac{S_{nk}^*(t,0)}{\sum_{i=1}^{n} c_{nik}} \right) \right\} m(t) dt \\
= \left( \sum_{j=1}^{n} c_{njk} \right) \int_0^1 g'(\varepsilon_t) [S_{nk}^*(t,y) - S_{nk}^*(t,0)] m(t) dt \\
= \left( \sum_{j=1}^{n} c_{njk} \right) \int_0^1 [g'(\varepsilon_t) - g'(t)] [S_{nk}^*(t,y) - S_{nk}^*(t,0)] m(t) dt + \\
+ \left( \sum_{j=1}^{n} c_{njk} \right) \int_0^1 g'(t) \left\{ [S_{nk}^*(t,y) - S_{nk}^*(t,0)] - [S_{nk}(t,y) - S_{nk}(t,0)] \right\} m(t) dt + \\
+ \left( \sum_{j=1}^{n} c_{njk} \right) \int_0^1 g'(t) [S_{nk}(t,y) - S_{nk}(t,0)] m(t) dt,
\]

...
\[ = I_y + II_y + III_y, \quad (4.4.6) \]

where \( \xi_t \) is between \( \frac{S_{nk}^*(t,y)}{\sum_{i=1}^n c_{nik}} \) and \( \frac{S_{nk}^*(t,0)}{\sum_{i=1}^n c_{nik}} \).

By Corollary 3.2.5, for any \( \epsilon > 0 \), there exists a positive integer \( N \) such that

\[
P\left( \sup \left\{ \frac{S_{nk}^*(t,y)}{\sum_{i=1}^n c_{nik}} - t \mid t \in [0,1], |y| \leq K \right\} > \delta \right) \leq \epsilon \quad \text{for } n \geq N,
\]

where \( \delta > 0 \) such that \( a < (a_o - \delta) < (b_o + \delta) < b \). Since \( g' \) is differentiable with a bounded derivative in \( [a_o - \delta, b_o + \delta] \), for \( n \geq N \), we have

\[ I_y = \left( \sum_{j=1}^n c_{njk} \right) \int_0^1 g''(\eta_t)(\xi_t - t)[S_{nk}^*(t,y) - S_{nk}^*(t,0)] m(t) dt \quad (4.4.7) \]

where \( \eta_t \) is between \( \xi_t \) and \( t \), and for any \( \rho > 0 \),

\[
P\left\{ \sup_{|y| \leq K} |I_y| > \rho \right\} \leq \epsilon + P\left\{ \sup_{|y| \leq K} |I_y| > \rho, \sup_{|y| \leq K} \left| \frac{S_{nk}^*(t,y)}{\sum_{i=1}^n c_{nik}} - t \right| < \delta \right\}.
\]

\[ (4.4.8) \]

When

\[ \left| \frac{S_{nk}^*(t,y)}{\sum_{i=1}^n c_{nik}} - t \right| \leq \delta, \]

we have \( \xi_t \in [a_o - \delta, b_o + \delta] \) for \( t \in [a_o, b_o] \), therefore \( \eta_t \in [a_o - \delta, b_o + \delta] \) for \( t \in [a_o, b_o] \), and, since \( \xi_t \sum_{i=1}^n c_{nik} \) is between \( S_{nk}^*(t,y) \) and \( S_{nk}^*(t,0) \),

\[ \sum_{i=1}^n c_{nik}(\xi_t - t) = [\xi_t \sum_{i=1}^n c_{nik} - S_{nk}^*(t,0)] \]
is between \([S_{nk}^*(t,y) - S_{nk}(t,0)]\) and \([S_{nk}^*(t,0) - S_{nk}(t,0)]\), therefore, by Corollary 3.2.5, we have

\[
\sup \left\{ \left| \sum_{i=1}^{n} c_{nik}(\xi_{t}-t) \right| ; t \in [0,1], |u| \leq K \right\} = O_p(1). \tag{4.4.9}
\]

Therefore,

\[
\sup_{|u| \leq K} |I_u| = O_p(1) \tag{4.4.10}
\]

follows from (4.4.7) through (4.4.9) and Corollary 3.2.2.

We notice that, for any \(|u| \leq K\), \(E\{I_u\} = 0\) and \(I_{\overline{u}} = 0\). Since

\[
I_u = \left( \sum_{j=1}^{n} c_{njk} \right) \left\{ \sum_{i=1}^{n} c_{nik} \int_{F(Y_{i} - \xi_{n}^{T}u)}^{F(Y_{i} - \xi_{n}^{T}u)} g'(t)m(t)dt - \int_{-\infty}^{1} g'(t)[S_{nk}(t,u) - S_{nk}(t,0)]m(t)dt \right\},
\]

for any \(u\) and \(y\),

\[
\text{Var}\{I_u - I_y\} \leq \left( \sum_{j=1}^{n} c_{njk} \right)^2 \sum_{i=1}^{n} c_{nik}^2 \left\{ \int_{F(Y_{i} - \xi_{n}^{T}u)}^{F(Y_{i} - \xi_{n}^{T}y)} g'(t)m(t)dt \right\}^2
\]

\[
\leq M_1 \left( \sum_{j=1}^{n} c_{njk} \right)^2 \sum_{i=1}^{n} c_{nik}^2 E\{F(Y_{i} - \xi_{n}^{T}u) - F(Y_{i} - \xi_{n}^{T}y)\}^2
\]

\[
\leq M_2 \left( \sum_{j=1}^{n} c_{njk} \right)^2 a_n^2 |u - y|^2 \leq M_2 C^2 |u - y|^2
\]
where $M_1$ and $M_2$ are constants. By the similar argument of the proof of Theorem 3.3.2, we have, as $n \to \infty$

$$\sup_{|u| \leq K} |\Pi_{II}^u| \xrightarrow{P} 0. \quad (4.4.11)$$

Since, by (B1), (B2) and (4.2.19),

$$\Pi_{III}^u - \left( \sum_{j=1}^{n} c_{njk} \right) \sum_{i=1}^{n} c_{nik} \epsilon_{ni}^T u \gamma =$$

$$= \left( \sum_{j=1}^{n} c_{njk} \right) \int_{0}^{1} \{ g'(t)[S_{nk}(t,\nu) - S_{nk}(t,\varrho)] - \sum_{i=1}^{n} c_{nik} \epsilon_{ni}^T u \} m(t) dt$$

$$= \left( \sum_{j=1}^{n} c_{njk} \right) \sum_{i=1}^{n} c_{nik} \int_{0}^{1} g'(t)[F(F^{-1}(t) + \epsilon_{ni}^T u)] -$$

$$- F(F^{-1}(t)) - \epsilon_{ni}^T u F'(F^{-1}(t)) \} m(t) dt$$

$$= \left( \sum_{j=1}^{n} c_{njk} \right) \sum_{i=1}^{n} c_{nik} \epsilon_{ni}^T u \int_{0}^{1} g'(t)[F'(\zeta_t) - F'(F^{-1}(t))] m(t) dt$$

$$\leq M_3 \left( \sum_{j=1}^{n} c_{njk} \right) \sum_{i=1}^{n} c_{nik} \epsilon_{ni}^T u \int_{0}^{1} |\zeta_t - F^{-1}(t)| \ |g'(t) m(t)| dt$$

$$\leq M_4 \left( \sum_{j=1}^{n} c_{njk} \right) a_{n} \sum_{i=1}^{n} c_{nik} \epsilon_{ni}^T u \leq M_4 C \sum_{i=1}^{n} c_{nik} \epsilon_{ni}^T u$$

where $\zeta_t$ is between $F^{-1}(t)$ and $\{F^{-1}(t) + \epsilon_{ni}^T u\}$, and $M_3$ and $M_4$ are constants, therefore, as $n \to \infty$

$$\sup_{|u| \leq K} \left| \Pi_{III}^u - \left( \sum_{j=1}^{n} c_{njk} \right) \sum_{i=1}^{n} c_{nik} \epsilon_{ni}^T u \gamma \right| = O(1). \quad (4.4.12)$$
Therefore, (4.4.5) follows from (4.4.6) and (4.4.10) through (4.4.12).

In the general case of \(\{c_{njk}\}\), we have the following proposition.

**COROLLARY 4.4.2.** Under all the assumptions of Lemma 4.4.1, except replacing (A1) and (A2) by (A3), we have that, for any \(K > 0\) and \(1 \leq k \leq p\), as \(n \to \infty\)

\[
\sup_{|u| \leq K} \left| \left( \sum_{j=1}^{n} c^{+}_{njk} \right)^{2} \int_{0}^{1} \left\{ g\left( \frac{S^{+}_{nk}(t,u)}{\sum_{i=1}^{n} c^{+}_{nik}} \right) - g\left( \frac{S^{+}_{nk}(t,0)}{\sum_{i=1}^{n} c^{+}_{nik}} \right) \right\} m(t) dt - \left( \sum_{j=1}^{n} c^{+}_{njk} \right) \sum_{i=1}^{n} c^{+}_{nik} c^{T}_{ni} u \gamma \right| = O_p(1),
\]

(4.4.13)

and

\[
\sup_{|u| \leq K} \left| \left( \sum_{j=1}^{n} c^{-}_{njk} \right)^{2} \int_{0}^{1} \left\{ g\left( \frac{S^{-}_{nk}(t,u)}{\sum_{i=1}^{n} c^{-}_{nik}} \right) - g\left( \frac{S^{-}_{nk}(t,0)}{\sum_{i=1}^{n} c^{-}_{nik}} \right) \right\} m(t) dt - \left( \sum_{j=1}^{n} c^{-}_{njk} \right) \sum_{i=1}^{n} c^{-}_{nik} c^{T}_{ni} u \gamma \right| = O_p(1).
\]

(4.4.14)

**Proof.** Consider any \(1 \leq k \leq p\). If \(d^{+}_{nk} = 0\) or \(d^{-}_{nk} = 0\), this is just the case of Lemma 4.4.1, and the left hand side of (4.4.13) or (4.4.14) is equal to 0. If \(d^{+}_{nk} > 0\) and \(d^{-}_{nk} > 0\), the proof is the same as Lemma 4.4.1's by using Corollary 3.2.3 and Corollary 3.2.6.
THEOREM 4.4.3. For a fixed \(k=1, \ldots, p\), assume (A1), (A2), (B1), (B2), (C1) and (C2). Then, for any \(K>0\), as \(n \to \infty\)

\[
\sup_{|y| \leq K} \left| \sum_{j=1}^{n} c_{nj} [M_{nj}(y) - M_{nj}(0)] + \sum_{i=1}^{n} c_{nki} x_{i} y \right| = O_{p}(1). 
\] (4.4.15)

Proof. Consider a functional \(\tau: D[0,1] \to \mathbb{R}\) defined by

\[
\tau(G) = \int_{0}^{1} g(G(t)) m(t) dt, \quad G \in D[0,1]. 
\] (4.4.16)

where \(g\) is given by (4.4.1). Then, \(\tau\) can be expressed as a composition of the following second-order Hadamard differentiable transformations:

\[
\gamma_{1}: D[0,1] \to L_{1}[0,1], \quad \gamma_{1}(S) = g \circ S;
\]

\[
\gamma_{2}: L_{1}[0,1] \to \mathbb{R}, \quad \gamma_{2}(S) = \int_{0}^{1} S(t) m(t) dt;
\]

and

\[
\tau(G) = \gamma_{2}(\gamma_{1}(G)).
\]

Note that \(\gamma_{1}\), by Proposition 4.3.2, is second-order Hadamard differentiable at \(U\) with

\[
\gamma'_{1}(H) = (g' \circ U) H, \quad \gamma''_{1}(H) = (g'' \circ U) H^2, \quad H \in D[0,1]
\]
because $g$ is differentiable everywhere and $g'$ is bounded, continuous and piecewise differentiable with a bounded derivative, and that $\gamma_2$ is linear and continuous, thus is second-order Hadamard differentiable with $\gamma''_2(G) = 0$. From Proposition 4.3.1, $\tau$ is second-order Hadamard differentiable at $U$ with

$$\tau'_{U}(G) = \int_0^1 g'(t) G(t)m(t)dt, \quad G \in D[0,1] \quad (4.4.17)$$

and

$$\tau''_{U}(G) = \int_0^1 g''(t) G^2(t)m(t)dt, \quad G \in D[0,1]. \quad (4.4.18)$$

From Theorem 4.3.3, we have, for any $K > 0$, as $n \to \infty$

$$\sup_{|y| \leq K} \left| \left( \sum_{i=1}^{n} \frac{c_{njk}}{\sum_{i=1}^{n} c_{nik}} \right)^2 \left\{ \tau_{\left( \sum_{i=1}^{n} c_{nik} \right)} - \tau(U(\cdot)) \right\} - \sum_{i=1}^{n} c_{njk} \tau'_{U}(S^*_{nk}(\cdot, y) - U(\cdot)) \sum_{i=1}^{n} c_{nik} \right| \overset{P}{\to} 0,$$

therefore, as $n \to \infty$

$$\sup_{|y| \leq K} \left| \left( \sum_{i=1}^{n} \frac{c_{njk}}{\sum_{i=1}^{n} c_{nik}} \right)^2 \left\{ \tau_{\left( \sum_{i=1}^{n} c_{nik} \right)} - \tau(S^*_{nk}(\cdot, 0)) \right\} - \sum_{i=1}^{n} c_{njk} \tau'_{U}(S^*_{nk}(\cdot, y) - S^*_{nk}(\cdot, 0)) \right| \overset{P}{\to} 0.$$

$$- \frac{1}{2} \left\{ \tau'''_{U}(S^*_{nk}(\cdot, y) - U(\cdot)) \sum_{i=1}^{n} c_{nik} \right\} - \tau''_{U}(S^*_{nk}(\cdot, 0) - U(\cdot)) \sum_{i=1}^{n} c_{nik} \right| \overset{P}{\to} 0. \quad (4.4.19)$$

Since $\psi$ is absolutely continuous,
\[ \tau'_u(S^*_{nk}(\cdot, y)) = \int_0^t g'(t) S^*_{nk}(t, y)m(t)dt \]
\[ = \sum_{i=1}^{\infty} c_{nik} \int_{\gamma_i - \epsilon_{n_i}^T y}^{1} \psi'(F^{-1}(t)) \frac{dt}{F'(F^{-1}(t))} = \sum_{i=1}^{n} c_{nik} \int_{\gamma_i - \epsilon_{n_i}^T y}^{\psi(\infty)} dx \]
\[ = \psi(\infty) \sum_{i=1}^{n} c_{nik} - M_{nk}(y), \]

then,
\[ \tau'_u(S^*_{nk}(\cdot, y) - S^*_{nk}(\cdot, 0)) = -[M_{nk}(y) - M_{nk}(0)]. \quad (4.4.20) \]

Since
\[ \tau''_u(S^*_{nk}(\cdot, y) - U(\cdot) \sum_{i=1}^{n} c_{nik}) = \int_0^1 g''(t)[S^*_{nk}(t, y) - t \sum_{i=1}^{n} c_{nik}]^2 m(t)dt \]
\[ = \int_0^1 g''(t)[S^*_{nk}(t, y) - S_{nk}(t, 0)]^2 m(t)dt, \]

by Corollary 3.2.5 and the boundedness of \( g''(t) \) for \( t \in [a_0, b_0] \), as \( n \to \infty \)

\[ \sup_{|y| \leq K} \left| \tau''_u(S^*_{nk}(\cdot, y) - U(\cdot) \sum_{i=1}^{n} c_{nik}) \right| = O_p(1). \quad (4.4.21) \]

Therefore, (4.4.15) follows from (4.4.19) through (4.4.21) and Lemma 4.4.1.

\[ \square \]

In the general case of \( \{c_{nik}\}, 1 \leq i \leq n, 1 \leq k \leq p, \) we have the following
THEOREM 4.4.4. Under all the assumptions of Theorem 4.4.3, except replacing (A1) and (A2) by (A3), assume

$$|B_n^+ B_n^-| = O(1), \quad \text{as } n \to \infty$$ (4.4.22)

where $B_n^+ = \text{Diag}(b_{n1}^+, \ldots, b_{np}^+)$, and

$$\bar{b}_n = \left( \frac{1}{b_{n1}^-} I\{b_{n1}^- > 0\}, \ldots, \frac{1}{b_{np}^-} I\{b_{np}^- > 0\} \right)^\top. $$ (4.4.23)

Then, for any $K > 0$ and $1 \leq k \leq p$, as $n \to \infty$

$$\sup_{|u| \leq K} \left| B_{nk}^+ [M_{nk}(u) - M_{nk}(0) + \sum_{i=1}^{p} c_{ni} c_{ni}^\top u \gamma] \right| = O_p(1),$$ (4.4.24)

where $B_{nk}^+ = b_{nk}^+ I\{b_{nk}^+ > 0\} + b_{nk}^- I\{b_{nk}^- = 0\}$. Therefore, as $n \to \infty$

$$\sup_{|u| \leq K} \left| \bar{B}_n^+ [M_n(u) - M_n(0) + Q_n u \gamma] \right| = O_p(1),$$ (4.4.25)

where $B_n^+ = \text{Diag}(\bar{b}_{n1}^+, \ldots, \bar{b}_{np}^+)$. 

Proof. Consider any $1 \leq k \leq p$. If $d_{nk}^+ = 0$ or $d_{nk}^- = 0$, (4.4.24) is the case of Theorem 4.4.3. If $d_{nk}^+ > 0$ and $d_{nk}^- > 0$, using the functional $\tau$ given by (4.4.16), from Theorem 4.3.4, we have, for any $K > 0$ and $1 \leq k \leq p$, as $n \to \infty$.
\[
\sup_{|y| \leq K} \left| \left( \sum_{i=1}^{n} c_{nik}^+ \right)^2 \left\{ r \left( \frac{S_{nk}^{*+}(\cdot, y)}{\sum_{i=1}^{n} c_{nik}^+} \right) - r \left( \frac{S_{nk}^{*+}(\cdot, 0)}{\sum_{i=1}^{n} c_{nik}^+} \right) \right\} - \sum_{i=1}^{n} c_{nik}^+ r_{U} (S_{nk}^{*+}(\cdot, y) - S_{nk}^{*+}(\cdot, 0)) - \right.
\]

\[- \frac{1}{2} \left\{ r_{U} (S_{nk}^{*+}(\cdot, y) - U(\cdot) \sum_{i=1}^{n} c_{nik}^+) - r_{U} (S_{nk}^{*+}(\cdot, 0) - U(\cdot) \sum_{i=1}^{n} c_{nik}^+) \right\} \overset{P}{\rightarrow} 0, \quad (4.4.26)\]

and

\[
\sup_{|y| \leq K} \left| \left( \sum_{i=1}^{n} c_{nik}^- \right)^2 \left\{ r \left( \frac{S_{nk}^{*-}(\cdot, y)}{\sum_{i=1}^{n} c_{nik}^-} \right) - r \left( \frac{S_{nk}^{*-}(\cdot, 0)}{\sum_{i=1}^{n} c_{nik}^-} \right) \right\} - \sum_{i=1}^{n} c_{nik}^- r_{U} (S_{nk}^{*-}(\cdot, y) - S_{nk}^{*-}(\cdot, 0)) - \right.
\]

\[- \frac{1}{2} \left\{ r_{U} (S_{nk}^{*-}(\cdot, y) - U(\cdot) \sum_{i=1}^{n} c_{nik}^-) - r_{U} (S_{nk}^{*-}(\cdot, 0) - U(\cdot) \sum_{i=1}^{n} c_{nik}^-) \right\} \overset{P}{\rightarrow} 0. \quad (4.4.27)\]

Since

\[
r_{U} (S_{nk}^{*+}(\cdot, y) - U(\cdot) \sum_{i=1}^{n} c_{nik}^+) = \int_{0}^{1} g''(t)[S_{nk}^{*+}(t, y) - t \sum_{i=1}^{n} c_{nik}^+]^2 m(t) dt
\]

\[= \int_{0}^{1} g''(t)[S_{nk}^{*+}(t, y) - S_{nk}^{*+}(t, 0)]^2 m(t) dt,
\]

and

\[
r_{U} (S_{nk}^{*-}(\cdot, y) - U(\cdot) \sum_{i=1}^{n} c_{nik}^-) = \int_{0}^{1} g''(t)[S_{nk}^{*-}(t, y) - t \sum_{i=1}^{n} c_{nik}^-]^2 m(t) dt
\]

\[= \int_{0}^{1} g''(t)[S_{nk}^{*-}(t, y) - S_{nk}^{*-}(t, 0)]^2 m(t) dt,
\]
by Corollary 3.2.6 and the boundedness of \( g'(t) \) for \( t \in [a_0, b_0] \), we have, as \( n \to \infty \)

\[
\sup_{|y| \leq K} \left| r''(S_{nk}^*(\cdot, y) - U(\cdot) \sum_{i=1}^{n} c_{nik}^+) \right| = O_p(1),
\]

(4.4.28)

and

\[
\sup_{|y| \leq K} \left| r''(S_{nk}^*(\cdot, y) - U(\cdot) \sum_{i=1}^{n} c_{nik}^-) \right| = O_p(1).
\]

(4.4.29)

Since

\[
r''(S_{nk}^*(\cdot, y)) = \int_0^1 g'(t) S_{nk}^*(t, y)m(t) dt
\]

\[
= \sum_{i=1}^{\infty} c_{nik}^+ \int_{F(Y_i - \xi_{n,i}^T y)}^1 \frac{\psi'(F^{-1}(t))}{F'(F^{-1}(t))} dt
= \sum_{i=1}^{n} c_{nik}^+ \int_{(Y_i - \xi_{n,i}^T y)}^{\infty} \psi(x) dx
\]

\[
= \psi(\infty) \sum_{i=1}^{n} c_{nik}^+ - M_{nk}^+(y),
\]

(4.4.30)

and

\[
r''(S_{nk}^*(\cdot, y)) = \int_0^1 g'(t) S_{nk}^*(t, y)m(t) dt
\]

\[
= \psi(\infty) \sum_{i=1}^{n} c_{nik}^- - M_{nk}^-(y),
\]

(4.4.31)

where

\[
M_{nk}^+(y) = \sum_{i=1}^{n} c_{nik}^+ \psi(Y_i - \xi_{n,i}^T y), \quad M_{nk}^-(y) = \sum_{i=1}^{n} c_{nik}^- \psi(Y_i - \xi_{n,i}^T y).
\]
so that

\[ M_{nk}(u) = M^+_{nk}(u) - M^-_{nk}(u), \quad (4.4.32) \]

then, as \( n \to \infty \)

\[ \sup_{|y| \leq K} \left| b^+_{nk}[M^+_{nk}(u) - M^+_{nk}(0) + \frac{N}{n} \sum_{i=1}^{n} c_i \xi_{ni}^T u \gamma] \right| = o_p(1) \quad (4.4.33) \]

and

\[ \sup_{|y| \leq K} \left| b^-_{nk}[M^-_{nk}(u) - M^-_{nk}(0) + \frac{N}{n} \sum_{i=1}^{n} c_i \xi_{ni}^T u \gamma] \right| = o_p(1) \quad (4.4.34) \]

follow from (4.4.13), (4.4.26), (4.4.28), (4.4.30) and (4.4.14), (4.4.27), (4.4.29), (4.4.31), respectively. Therefore, as \( n \to \infty \)

\[
\begin{align*}
\sup_{|y| \leq K} & \left| b^+_{nk}[M^+_{nk}(u) - M^+_{nk}(0) + \frac{N}{n} \sum_{i=1}^{n} c_i \xi_{ni}^T u \gamma] + \\
+b^-_{nk}[M^-_{nk}(u) - M^-_{nk}(0) + \frac{N}{n} \sum_{i=1}^{n} c_i \xi_{ni}^T u \gamma] \right| \\
= & \sup_{|y| \leq K} \left| b^+_{nk}[M^+_{nk}(u) - M^+_{nk}(0) + \frac{N}{n} \sum_{i=1}^{n} c_i \xi_{ni}^T u \gamma] + \\
+(b^-_{nk} - b^+_{nk})[M^-_{nk}(u) - M^-_{nk}(0) + \frac{N}{n} \sum_{i=1}^{n} c_i \xi_{ni}^T u \gamma] \right| \\
= & \sup_{|y| \leq K} \left| b^+_{nk}[M^+_{nk}(u) - M^+_{nk}(0) + \frac{N}{n} \sum_{i=1}^{n} c_i \xi_{ni}^T u \gamma] + \\
+(b^-_{nk} - 1)b^+_{nk}[M^-_{nk}(u) - M^-_{nk}(0) + \frac{N}{n} \sum_{i=1}^{n} c_i \xi_{ni}^T u \gamma] \right| = o_p(1).
\end{align*}
\]

Hence, (4.4.24) follows from (4.4.22) and (4.4.33) through (4.4.35).

\[ \Box \]

COROLLARY 4.4.5. Under all the assumptions of Theorem 4.4.3,
except replacing (A1) and (A2) by (A3), assume

\[ |B_\hat{n} \hat{F}_n^+| = O(1), \quad \text{as } n \to \infty \]  \hspace{1cm} (4.4.36)

where \( B_\hat{n} = \text{Diag}(b_{n1}, \ldots, b_{np}) \), and

\[ \hat{b}_n^+ = \left( \frac{1}{b_{n1}^+} \hat{I}\{b_{n1}^+ > 0\}, \ldots, \frac{1}{b_{np}^+} \hat{I}\{b_{np}^+ > 0\} \right)^T. \]  \hspace{1cm} (4.4.37)

Then, for any \( K > 0 \) and \( 1 \leq k \leq p \), as \( n \to \infty \)

\[ \sup_{|u| \leq K} \left| \hat{b}_{nk}^-[M_{nk}(u) - M_{nk}(0) + \sum_{i=1}^n c_{nik} e_{ni}^T u \gamma)] = O_p(1), \right. \]  \hspace{1cm} (4.4.38)

where \( \hat{b}_{nk}^- = b_{nk}^{-} \hat{I}\{b_{nk}^- > 0\} + b_{nk}^+ \hat{I}\{b_{nk}^- = 0\} \). Therefore, as \( n \to \infty \)

\[ \sup_{|u| \leq K} \left| \hat{B}_n [M_n(u) - M_n(0) + Q_n u \gamma)] = O_p(1), \right. \]  \hspace{1cm} (4.4.39)

where \( \hat{B}_n = \text{Diag} \{ \hat{b}_{n1}, \ldots, \hat{b}_{np} \} \).

Proof. The proof is similar to the one of Theorem 4.4.4. \( \square \)

The following theorem gives an asymptotic representation of M-estimators of regression.
THEOREM 4.4.6. In addition to (A3), (A4), (B1), (B2) and (C1) through (C4), assume that (4.4.22) and (4.4.36) hold. Then, as \( n \to \infty \)

\[
C_n^2(\hat{\beta}_n - \beta) = \frac{1}{4} Q_n^{-1} M_n(0) + O_p \left( \frac{1}{(I^T B_n^+)^{1/2}} \right),
\]

(4.4.40)

where \((x) = ((x_1, \ldots, x_q)) = \min_{1 \leq i \leq q} |x_i|\) and \(I = (1, \ldots, 1)^T\).

Proof. From the assumptions on \( \psi \), we have that, for arbitrary \( \epsilon > 0 \), \( 1 \leq k \leq p \) and \( n \geq 1 \), there exists \( K_0 > 0 \) such that

\[
P\{|M_{nk}(0)| > K_0\} \leq \frac{\text{Var}\{M_{nk}(0)\}}{K_0^2} = \frac{\sigma^2}{K_0^2} < \epsilon
\]

where \( \sigma^2 = \int \psi^2 dF < \infty \), therefore,

\[
|M_n(0)| = O_p(1).
\]

(4.4.41)

Since (A3), (4.4.22) and (4.4.36) imply

\[
(I^T B_n^+) \to \infty, \quad \text{as} \quad n \to \infty
\]

(4.4.42)

by Theorem 4.4.4, we have, for any \( K > 0 \), as \( n \to \infty \)

\[
\sup_{|u| \leq K} |M_n(u) - M_n(0) + Q_n u\gamma| \overset{P}{\to} 0.
\]

(4.4.43)
Using the similar argument referring to (3.3.16) in the proof of (2) of Lemma 3.3.1, by (4.4.1), (4.4.43), (A4) and (C3), we have that, as \( n \to \infty \)

\[
|\hat{u}_n| = O_p(1).
\]  

(4.4.44)

Therefore, by Theorem 4.4.4,

\[
\mathcal{B}_n^+(Q_n \hat{u}_n \gamma - M_n(\theta)) = O_p(1),
\]  

(4.4.45)

as \( n \to \infty \). Since (4.4.45) implies

\[
Q_n \hat{u}_n \gamma = M_n(\theta) + O_p\left(\frac{1}{(1 + \mathcal{B}_n^+)}\right),
\]  

(4.4.46)

hence, (4.4.40) follows from (4.4.46) and (A4). \( \square \)
NOTATIONS AND ABBREVIATIONS

\( x, y, u, v, \ldots, A, B, C, \ldots \): column vectors or matrices

\( I: = (1, \ldots, 1)^T \)

\( I(x \geq y): = 1 \) if \( x \geq y \) and \( = 0 \) if \( x < y \)

\( \| \cdot \| \): standard Euclidean norm for any vector \( x \in \mathbb{R}^q \), where \( q \) is a positive integer; uniform norm for any \( H \in \mathbb{D}[0,1] \).

\( \| \cdot \|_{L_p} \): standard norm in \( L^P \) space

\( \cdot \): \( \| x \| = \max_{1 \leq i \leq q} |x_i| \) for \( x = (x_1, \ldots, x_q) \in \mathbb{R}^q \), where \( q \) is a positive integer

\( (\cdot) \): \( \min_{1 \leq i \leq q} |x_i| \) for \( x = (x_1, \ldots, x_q) \in \mathbb{R}^q \), where \( q \) is a positive integer

\( \lambda_1(A) \): smallest eigenvalue of matrix \( A \)

\( \lambda_2(A) \): largest eigenvalue of matrix \( A \)

\( E_q \): space \([0,1]^q\), where \( q \) is a positive integer

\( L_q(m) = \{ \frac{1}{m}(l_1, \ldots, l_q) \mid l_i = 0, 1, \ldots, m; i = 1, \ldots, q \} \), where \( q \) and \( m \) are positive integers

\( \lim \): lim sup

\( \lim \): lim inf
\(D\): convergence in distribution

\(P\): convergence in probability

\(N_p(\theta, \Sigma)\): p-variate normal distribution with mean \(\theta\) and covariance matrix \(\Sigma\)

\(o(1)\): term tends to 0

\(O(1)\): term tends to be bounded

a.e.: almost everywhere

a.s.: almost sure

d.f.: distribution function

i.i.d.r.v.: independently and identically distributed random variables

sup: supremum

inf: infimum

LSE: least square estimator

MLE: maximum likelihood estimator

DCT: dominate convergence theorem
REFERENCES


