UN
SECURI

AD-A219 921

IENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY N/A	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for Public Release; Distribution Unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) Technical Report No. 283	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 90-0326
6a. NAME OF PERFORMING ORGANIZATION University of North Carolina Center for Stochastic Processes 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION AFOSR/NM
6c. ADDRESS (City, State, and ZIP Code) Statistics Department CB #3260, Phillips Hall Chapel Hill, NC 27599-3260	7b. ADDRESS (City, State, and ZIP Code) Bldg. 410 Bolling Air Force Base, DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR 8b. OFFICE SYMBOL (If applicable) NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F49620 85C 0144
8c. ADDRESS (City, State, and ZIP Code) Bldg. 410 Bolling AFB, DC 20332-6448	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 6.1102F PROJECT NO. 2304 TASK NO. A5 WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
M-estimators in linear models with long range dependent errors

12. PERSONAL AUTHOR(S)
Koul, H.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
preprint	FROM ______ TO ______	1990, February	19

16. SUPPLEMENTARY NOTATION
N/A

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Hermite rank and polynomials, Linearity.
XXXXXX	XXXXXXXXXX	XXX	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This note discusses the asymptotic behavior of a class of M-estimators in linear models when errors are Gaussian, or a function of Gaussian random variables, that are long range dependent. The asymptotics are discussed when the design variables are either i.i.d. or long range dependent, independent of the errors, or known constants. It is observed that the class of M-estimators of the regression parameter vector corresponding to skew symmetric scores and symmetric errors asymptotically behave like the least squares estimators. Moreover, in these cases, if the design variables are either i.i.d. or known constants then the limiting distributions are Normal. But if the design variables are also long range dependent then the limiting distributions are nonnormal.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED	
22a. NAME OF RESPONSIBLE INDIVIDUAL Professor Eytan Barouch	22b. TELEPHONE (Include Area Code) (202)767-5026	22c. OFFICE SYMBOL AFOSR/NM

DD Form 1473, JUN 86 Previous editions are obsolete. SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFED

DTIC ELECTE MAR 29 1990

CENTER FOR STOCHASTIC PROCESSES

Department of Statistics
University of North Carolina
Chapel Hill, North Carolina

M-Estimators in Linear Models With
Long Range Dependent Errors

by

Hira L. Koul

February 1990

Technical Report No. 283

219. J. Bather. Stopping rules and ordered families of distributions. Dec. 87. *Sequential Anal.*, 7, 1988, 111-126.

220. S. Cambanis and M. Maejima. Two classes of self-similar stable processes with stationary increments. Jan. 88. *Stochastic Proc. Appl.*, 32, 1989, 305-329.

221. H.P. Hucke, G. Kallianpur and R.L. Karandikar. Smoothness properties of the conditional expectation in finitely additive white noise filtering. Jan. 88. *J. Multivariate Anal.*, 27, 1988, 261-269.

222. I. Mitoma. Weak solution of the Langevin equation on a generalized functional space. Feb. 88. (Revised as No. 238).

223. L. de Haan, S.I. Resnick, H. Rootzén and C. de Vries. Extremal behaviour of solutions to a stochastic difference equation with applications to arch-processes. Feb. 88. *Stochastic Proc. Appl.*, 32, 1989, 213-224.

224. O. Kallenberg and J. Szulga. Multiple integration with respect to Poisson and Lévy processes. Feb. 88. *Prob. Theor. Rel. Fields*, 83, 1989, 101-134.

225. D.A. Dawson and L.G. Gorostiza. Generalized solutions of a class of nuclear space valued stochastic evolution equations. Feb. 88. *Appl. Math. Optimization*, to appear.

226. G. Samorodnitsky and J. Szulga. An asymptotic evaluation of the tail of a multiple symmetric α-stable integral. Feb. 88. *Ann. Probability*, 17, 1989, 1503-1520.

227. J.J. Hunter. The computation of stationary distributions of Markov chains through perturbations. Mar. 88.

228. H.C. Ho and T.C. Sun. Limiting distribution of nonlinear vector functions of stationary Gaussian processes. Mar. 88. *Ann. Probability*, to appear.

229. R. Brigola. On functional estimates for ill-posed linear problems. Apr. 88.

230. M.R. Leadbetter and S. Nandagopalan. On exceedance point processes for stationary sequences under mild oscillation restrictions. Apr. 88. *Proc. Oberwolfach Conf. on Extremal Value Theory*, J. Hüsler and R. Reiss, eds., Springer, 1989, 69-80.

231. S. Cambanis, J. P. Nolan and J. Rosinski. On the oscillation of infinitely divisible processes. Apr. 88. *Stochastic Proc. Appl.*, to appear.

232. G. Hardy, G. Kallianpur and S. Ramasubramanian. A nuclear space-valued stochastic differential equation driven by Poisson random measures. Apr. 88.

233. D.J. Daley, T. Rolski. Light traffic approximations in queues (II). May 88. *Math. Operat. Res.*, to appear.

234. G. Kallianpur, I. Mitoma, R.L. Wolpert. Nuclear space-valued diffusion equations. July 88. *Stochastics*, 1989, to appear.

235. S. Cambanis. Admissible translates of stable processes: A survey and some new models. July 88.

236. E. Platen. On a wide range exclusion process in random medium with local jump intensity. Aug. 88.

237. R.L. Smith. A counterexample concerning the extremal index. Aug. 88. *Adv. Appl. Probab.*, 20, 1988, 681-683.

238. G. Kallianpur and I. Mitoma. A Langevin-type stochastic differential equation on a space of generalized functionals. Aug. 88.

239. C. Houdré. Harmonizability, V-boundedness, (2,P)-boundedness of stochastic processes. Aug. 88. *Prob. Th. Rel. Fields*, to appear.

240. G.W. Johnson and G. Kallianpur. Some remarks on Hu and Meyer's paper and infinite dimensional calculus on finitely additive canonical Hilbert space. Sept. 88. *Th. Prob. Appl.*, to appear.

241. L. de Haan. A Brownian bridge connected with extreme values. Sept. 88. *Sankhya*, 1989, to appear.

242. O. Kallenberg. Exchangeable random measures in the plane. Sept. 88. *J. Theor. Probab.*, to appear.

243. E. Masry and S. Cambanis. Trapezoidal Monte Carlo integration. Sept. 88. *SIAM J. Numer. Anal.*, 1989, to appear.

244. L. Pitt. On a problem of H.P. McKean. Sept. 88. *Ann. Probability*, 17, 1989, 1651-1657.

245. C. Houdré. On the linear prediction of multivariate (2,P)-bounded processes. Sept. 88.

246. C. Houdré. Stochastic processes as Fourier integrals and dilation of vector measures. Sept. 88. *Bull. Amer. Math. Soc.*, to appear.

247. J. Mijnheer. On the rate of convergence in Strassen's functional law of the iterated logarithm. Sept. 88. *Probab. Theor. Rel. Fields*, to appear.

248. G. Kallianpur and V. Perez-Abreu. Weak convergence of solutions of stochastic evolution equations on nuclear spaces. Oct. 88. Proc. Trento Conf. on Infinite Dimensional Stochastic Differential Equations, 1989, to appear.

249. R.L. Smith. Bias and variance approximations for estimators of extreme quantiles. Nov. 88.

250. H. Hurd. Spectral coherence of nonstationary and transient stochastic processes. Nov. 88. 4th Annual ASSP Workshop on Spectrum Estimation and Modeling, Minneapolis, 1988, 387-390.

251. J. Leskow. Maximum likelihood estimator for almost periodic stochastic processes models. Dec. 88.

252. M.R. Leadbetter and T. Hsing. Limit theorems for strongly mixing stationary random measures. Jan. 89. *Stochastic Proc. Appl.*, to appear.

253. M.R. Leadbetter, I. Weissman, L. de Haan, H. Rootzén. On clustering of high values in statistically stationary series. Jan. 89. *Proc. 4th Int. Meeting on Stat. Climatology*, to appear.

254. J. Leskow. Least squares estimation in almost periodic point processes models. Feb. 89.

255. N.N. Vakhania. Orthogonal random vectors and the Hurwitz-Radon-Eckmann theorem. Apr. 89.

M – Estimators in Linear Models With Long Range Dependent Errors

by

Hira L. Koul

Department of Statistics and Probability
Michigan State University
East Lansing, MI 48824

and

Center for Stochastic Processes
Department of Statistics
University of North Carolina
Chapel Hill, NC 27599–3260

Abstract. This note discusses the asymptotic behavior of a class of M – estimators in linear models when errors are Gaussian, or a function of Gaussian random variables, that are long range dependent. The asymptotics are discussed when the design variables are either i.i.d. or long range dependent, independent of the errors, or known constants. It is observed that the class of M – estimators of the regression parameter vector corresponding to skew symmetric scores and symmetric errors asymptotically behave like the least squares estimators. Moreover, in these cases, if the design variables are either i.i.d. or known constants then the limiting distributions are Normal. But if the design variables are also long range dependent then the limiting distributions are nonnormal.

AMS 1980 Subject Classification: Primary 62M10, 65G30.

Key words and phrases: Hermite rank and polynomials. Linearity.

Part of research supported by the Air Force Office of Scientific Research Contract No. F49620 85C 0144.

1. ***Introduction and Summary.*** A discrete time stationary stochastic process is said to be long range dependent if its covariances decrease to zero like a power of lag as the lag tends to infinity but their absolute sum diverges. Such processes arise in applications in Hydrology, Economics, Time Series Analysis and other sciences. See, e.g., the review paper by Mandelbrot and Taqqu (1979) and references therein for the importance of these processes. See Granger and Joyeux (1980), and Hosking (1981) for the usefulness of these processes in Economics and Time Series Analysis. For many technical results on these processes, see Taqqu (1975, 1979), Fox and Taqqu (1987) and Dehling and Taqqu (1989), and Yajima (1985, 1988), among others.

One of the popular class of estimators in linear models that has evolved over the last two and a half decades is the so called class of M – estimators. Most of the asymptotic literature on these estimators assumes either independent errors (Huber: 1981 and references therein) or weakly dependent errors, like strongly mixing, as in Koul (1977).

Because of the importance of both, M – estimators and the long range dependence, it is of interest to study the large sample behavior of these estimators in a linear regression setting when errors are either long range dependent Gaussian or functions of such random variables (r.v.'s). About the design variables in the linear model we shall assume that they are either r.v.'s or known constants. In the former case it will be further assumed that the design variables are independent of the errors and either i.i.d. or long range dependent. The case of the known constant designs will be discussed in Section 3. We shall for the time being restrict our attention to the case of random designs.

Accordingly, let $\eta_1, \eta_2, \ldots$ be a sequence of strictly stationary mean zero unit variance Gaussian r.v.'s with $\rho(k) := E\eta_1\eta_{k+1}$, $k \geq 0$. Let $\xi_1, \xi_2, \ldots$ be a sequence of observable p×1 stationary mean zero random vectors with $\Gamma(k) := E\xi_1\xi_{1+k}'$, $k \geq 0$. Consider the linear model

$$(1) \qquad Y_i = X_i'\beta + \epsilon_i, \qquad X_i' = (1, \xi_i'), \quad \beta \in \mathscr{R}^{p+1}, \ i \geq 1$$

where $\epsilon_i \equiv G(\eta_i)$, $i \geq 1$, G a measurable function form $\mathscr{R}$ to $\mathscr{R}$.

Note that the marginal distribution of ϵ_1 need not be Gaussian. In fact if one were to have a linear regression model with stationary errors whose marginal distribution function (d.f.) is F, then choosing $G = F^{-1}(\Phi)$ would yield the desired errors. Here Φ is the d.f. of a N(0,1) r.v. and $F^{-1}(u) = \inf\{x; F(x) \geq u\}$, $0 \leq u \leq 1$.

The class of M – estimators, one corresponding to each ψ, is defined as a solution $\hat{\beta}_N$ of the equation

(2) $$S(\mathbf{t}) := \sum_{i=1}^{N} \mathbf{X}_i \psi(Y_i - \mathbf{X}_i' \mathbf{t}) = 0$$

where ψ is a measurable function from $\mathscr{R}$ to $\mathscr{R}$ with

(3) $$E\psi(\epsilon) = 0, \ 0 < E\psi^2(\epsilon) < \infty.$$

Here, and in the sequel, η, ϵ, ξ etc. are copies of η_1, ϵ_1, ξ_1 etc. Also for a p×1 vector $\mathbf{t} \in \mathscr{R}^p$, $\mathbf{t}'$ will denote its transpose and $\|\mathbf{t}\|$ will stand for its Euclidean norm.

The present paper is concerned with investigating the large sample behavior of M – estimators when the r.v.'s $\{\eta_i\}$, in addition, satisfy

(4) $$\rho(k) = k^{-D_1} L_1(k), \quad 0 < D_1 < 1, k \geq 1$$

where $L_1(k)$ is positive for large k and slowly varying at infinity, i.e., $L_1(tx)/L_1(t) \to 1$ as $t \to \infty$ for every $x \in \mathscr{R}$.

About $\{\xi_i\}$ we shall additionally assume that

(5) $$\{\xi_i\} \text{ are independent of } \{\epsilon_i\}$$

and either

(6a) $$\xi_1, \xi_2, \ldots \quad \text{are i.i.d. r.v.'s}$$

or

(6b) $\{\xi_i\}$ are dependent with $\Gamma(k) = k^{-D_2}\mathscr{L}(k)$, $0 < D_2 < 1$,

where $\mathscr{L}$ is a p×p matrix of slowly varying functions at infinity and $\mathscr{L}(k)$ are positive definite for all large k.

The processes that have covariances like (4) or (6b) are called long range dependent. These covariances tend to zero but not fast enough so as to be summable.

In the case when errors are independent or weakly dependent, $A_N(\hat{\beta}_N - \beta)$ turns out to be asymptotically normally distributed where A_N equals $N^{\frac{1}{2}}$ in the case $\{\xi_i\}$ are i.i.d. r.v.'s or A_N equals $(X'X)^{\frac{1}{2}}$ in the case $\{\xi_i\}$ are the known constants. Here $X'X = \sum_{i=1}^{N} X_iX_i'$.

Recall that the way this result is proved is first to approximate $\hat{\beta}_N - \beta$ by $\{\sum_{i=1}^{N} X_iX_i' \psi'(\epsilon_i)\}^{-1}\cdot S(\beta)$. Then, by the LLN's, the first term in this approximation is seen to be of the order N^{-1} and this N^{-1} is split so as to stabilize $S(\beta)$ and $\hat{\beta}_N - \beta$. In the case the errors are independent or weakly dependent and the design variables are random, the scores $S(\beta)$ are of the order $O_p(N^{1/2})$ and hence one must have $A_N = N^{1/2}$. Note that, in view of the Ergodic Theorem, the first term in the above approximation is $O_p(N^{-1})$ as long as the summands $\{X_iX_i' \psi'(\epsilon_i)\}$ are stationary, ergodic, have finite first moments and $\{E[X_1X_1' \psi'(\epsilon_1)]\}^{-1}$ exists, regardless of whether the r.v.'s are long range dependent or not. Hence, even in the present case, the magnitude of $S(\beta)$ determines that of $\hat{\beta}_N - \beta$. The exposition in Section 2 below uses this observation. A similar observation is used in Section 3 when the design variables are the known constants.

One of the observations of this note is that the class of M – estimators corresponding to the skew symmetric scores and symmetric errors (i.e. skew symmetric G) asymptotically behaves like the least squares estimator under (6a) or (6b) or the known constant design case. This result, in the cases of (6a) and (6b), is stated and proved in Section 2 and in the other case, in Section 3, below. A similar observation was made by Beran and Kunsch

(1985) in connection with the one sample location model. We further observe that in these cases if the design variables are either i.i.d. or known constants then the limiting distributions are Normal. But if the design variables are also strongly dependent and there is no intercept parameter in the model then the the limiting distributions are nonnormal and appear at the end of Section 2.

In what follows, L, with or without suffix is a generic notation for a slowly varying function. All limits are taken as $N \to \infty$, unless mentioned otherwise. Also in most of our discussion the design variables need not be Gaussian.

2. ***The Case of Random – Designs.*** A preliminary result needed for obtaining a first order approximation to M – estimators is the asymptotic uniform linearity of S. The following theorem gives a set of sufficient conditions for such a result to hold. It also gives the required approximation to M – estimators. The statement of the Theorem is somewhat self contained.

***Theorem* 1.** *Let* $(\xi_1', \epsilon_1), (\xi_2', \epsilon_2)$ *be a strictly stationary sequence of random vectors with* ξ_i *being* $p \times 1$. *Let* $\mathbf{X}_i' = (1, \xi_i')$,

$$Y_i = \mathbf{X}_i'\boldsymbol{\beta} + \epsilon_i, \quad \textit{for some } \boldsymbol{\beta} \in \mathcal{R}^{p+1}, \ i \geq 1.$$

In addition assume the following:

(a) *The score function* ψ *satisfies* (1.3) *and is absolutely continuous with* a.e. *derivative* ψ' *satisfying* $E|\psi'| < \infty$, *and* ,

$$E\|\mathbf{X}_1\|^2 \, |\psi'(\epsilon - z\|\mathbf{X}_1\|) - \psi'(\epsilon)| \longrightarrow 0 \quad \textit{as } z \longrightarrow 0.$$

(b) *For* $N \geq p+1$, *there are sequences* $\{\mathbf{A}_N\}$ *and* $\{\mathbf{B}_N\}$ *of* $(p+1)\times(p+1)$ *matrices which are positive definite for sufficiently large* N *and satisfy*

(i) $\|\mathbf{B}_N^{-1}\| \longrightarrow 0, \ \|\mathbf{A}_N^{-1}\| \longrightarrow 0, \ N \cdot \|\mathbf{A}_N^{-1}\| \cdot \|\mathbf{B}_N^{-1}\| \longrightarrow 1$.

(ii) $\|\mathbf{B}_N^{-1} \cdot \mathbf{S}(\boldsymbol{\beta})\| = O_p(1)$, .

Then, for every $0 < b < \infty$,

$$(1) \qquad E \sup_{\|\Delta\| \le b} \| B_N^{-1}[S(\beta + A_N^{-1}\Delta) - S(\beta)] + B_N^{-1}\sum_i X_i X_i' \, \psi'(\epsilon_i) A_N^{-1}\Delta \| = o_p(1)$$

where S *is as in* (1.2).

In addition, if $\{\xi_i\}$ *are independent of* $\{\epsilon_i\}$ *and if* $E\|\xi\|^2 < \infty$, *then the random coefficient of* Δ *in the linear expansion* (1) *may be replaced by* $\mathbf{R} \cdot E\psi'(\epsilon)$ *where* $\mathbf{R} := E X_1 X_1'$.

Furthermore, if

(c) $\mathbf{R}^{-1}$ *exists, and* d) $0 < E\psi'(\epsilon)$,

then

$$(2) \qquad A_N(\hat{\beta}_N - \beta) = [\mathbf{R}\, E\psi'(\epsilon)]^{-1} \cdot B_N^{-1} S(\beta) + o_p(1).$$

Remark 1. It is perhaps worth repeating that in the above theorem neither $\{\xi_i\}$ nor $\{\epsilon_i\}$ need be Gaussian or functions of Gaussian r.v.'s.

Proof. From the definition of **S**, $S(\beta + A_N^{-1}\Delta) = \sum_i X_i \, \psi(\epsilon_i - X_i' A_N^{-1}\Delta)$. Now , use the definition of absolute continuity and routine arguments to get that the

$$\text{L.H.S.}(1) \le b\, E \sum_i \|B_N^{-1} X_i\| \cdot \|X_i\| \int_{-\|A_N^{-1}\|}^{\|A_N^{-1}\|} |\psi'(\epsilon - z\|X_i\|b) - \psi'(\epsilon)|\, dz$$

$$\le b\, N \cdot \|B_N^{-1}\| \cdot \|A_N^{-1}\| \cdot \int_{-\|A_N^{-1}\|}^{\|A_N^{-1}\|} E\|X_1\|^2 |\psi'(\epsilon - z\|X_1\|b) - \psi'(\epsilon)|\, dz \longrightarrow 0,$$

by (a) and (b)(i).

The claim about replacing $B_N^{-1}\sum_i X_i X_i' \psi'(\epsilon_i) A_N^{-1}$ by $\mathbf{R} \cdot E\psi'(\epsilon)$ follows from the Ergodic Theorem. The claim (2) is obtained from (1), (a)(ii), (c) and (d), with the help of Scheweder fixed point Theorem, just as in Huber (1981). □

***Remark* 2.** Observe that $\psi(x) \equiv x$ a priori satisfies (a). Another example of ψ satisfying (a) is the Huber function $\psi(x) := xI(|x|\le c) + c\,\mathrm{sgn}(x)$, $c > 0$, provided $\{\xi_i\}$ are independent of $\{\epsilon_i\}$, $E\|\xi\|^2 < \infty$, and F is continuous at $\pm c$. To see this observe that for this ψ the

$$\text{L.H.S.(a)} \le E\|\mathbf{X}_1\|^2\{[F(c+z\|\mathbf{X}_1\|) - F(c-z\|\mathbf{X}_1\|] + [F(-c+z\|\mathbf{X}_1\|) - F(-c-z\|\mathbf{X}_1\|)]\}.$$

Now the Dominated Convergence Theorem gives the claim. □

Observe that so far we have not used (1.4) or (1.6a) or (1.6b) or even the assumption about $\{\eta_i\}$ being Gaussian. We shall now use these assumptions to determine the sequences of matrices $\{\mathbf{A}_N\}$ and $\{\mathbf{B}_N\}$. The main requirement on $\mathbf{B}_N$ is (b)(ii). Once $\mathbf{B}_N$ is determined, $\mathbf{A}_N$ can be determined from (b)(i).

In order to assess the magnitude of $\mathbf{S}$ (write $\mathbf{S}$ for $\mathbf{S}(\beta)$) we shall use the Hermite expansion of $L_2(\mathcal{R}; d\Phi)$ functions. What follows about Hermite expansions etc. is borrowed from Feller (1971) and Taqqu (1975). With $\{H_q, q \ge 0\}$ denoting the Hermite polynomials, let $J_q := E\psi_1(\eta)H_q(\eta)$, where $\psi_1 = \psi(G)$. Let $m := \min\{q\ge 1, J_q \ne 0\}$ denote the Hermite rank of $\psi_1(\eta)$. The Hermite expansion of rank m of $\psi_1(\eta)$ is given by $\sum_{q=m}^{\infty} \frac{J_q}{q!} H_q(\eta)$.

Recall from Feller (1971) that $\{H_q(\eta_i)\}$ is a set of orthonormal r.v.'s in $L_2(\mathcal{R}; d\Phi)$ satisfying

$$H_0(x) \equiv 1,\ EH_q(\eta) = 0,\ q \ge 1; \tag{3}$$

$$EH_q(\eta_i)H_n(\eta_j) = \begin{cases} 0, & q \ne n \\ q!\rho^q(i-j), & q = n \end{cases},\ \forall\, i, j.$$

Now, we begin the argument for determining $\mathbf{B}_N$ and $\mathbf{A}_N$. For a $\boldsymbol{\lambda} \in \mathbb{R}^{p+1}$, write $\boldsymbol{\lambda}' = (\lambda_1, \boldsymbol{\lambda}_2')$, $\lambda_1 \in \mathcal{R}$, $\boldsymbol{\lambda}_2 \in \mathcal{R}^p$. From (1.5) and (3), $\forall\, \boldsymbol{\lambda} \in \mathcal{R}^{p+1}$,

$$E[\boldsymbol{\lambda}' \sum_i \mathbf{X}_i H_m(\eta_i)]^2 = m! \sum_{i=1}^{N} \sum_{j=1}^{N} [\lambda_1^2 + \boldsymbol{\lambda}_2' \,\boldsymbol{\Gamma}(i-j)\, \boldsymbol{\lambda}_2]\rho^m(i-j). \tag{4}$$

At this point we need to consider (1.6a) and (1.6b) separately.

Suppose **that (1.6a) holds**. Then the

$$\mathrm{RHS}(4) = m![\lambda_1^2 \sum_i \sum_j \rho^m(i-j) + \lambda_2'\lambda_2 N]$$

Now, if we restrict $D_1 < 1/m$, then from Taqqu (1975; Lemma 3) it follows that the

$$\mathrm{RHS}(4) \simeq c_1\lambda_1^2 N^{2-mD_1} L(N) + m!\, \lambda_2'\lambda_2 N.$$

where c_1 is a constant depending on D_1 and m. Thus in this case if we choose

$$\mathbf{B}_N = \begin{bmatrix} N^{H_1}L(N) & \mathbf{0}_{1\times p} \\ \mathbf{0}_{p\times 1} & N^{\frac{1}{2}}\mathbf{I}_{p\times p} \end{bmatrix} = \begin{bmatrix} b_{N1} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{N2} \end{bmatrix}, \text{ say,} \tag{5}$$

with $2H_1 = 2-mD_1$, then we see that

$$E(\lambda' \mathbf{B}_N^{-1} \sum_i X_i H_m(\eta_i))^2 = O(1) \ \ \forall\, \lambda \in \mathscr{R}^{p+1}. \tag{6}$$

We note that $D_1 < 1/m$ implies that $\{\psi(\epsilon_i)\}$ are also long range dependent. The case $D_1 \geq 1/m$ would yield that these r.v.'s are asymptotically weakly dependent and not interesting to us from the current point of view.

Now **suppose that (1.6b) holds**. Then the

$$\mathrm{RHS}(4) \approx m![\lambda_1^2 \sum_i \sum_j |i-j|^{-mD_1} L(i-j) + \sum_i \sum_j \lambda_2' \mathscr{L}(i-j)\, \lambda_2\, |i-j|^{-D_2-mD_1}]$$

Note that $\mathscr{L}$ being a matrix of slowly varying functions at infinity and that $\mathscr{L}(k)$ being positive definite for all large k, it follows that for every $\lambda_2 \in \mathscr{R}^p$, $\lambda_2'\mathscr{L}\lambda_2$ is slowly varying at infinity and that $\lambda_2'\mathscr{L}(k)\,\lambda_2 > 0$ for all large k and for every $\lambda_2 \in \mathscr{R}^p$.

Once again, use the arguments as in Taqqu (op cit.) to conclude that the

$$\mathrm{R.H.S.}(4) \simeq c_1\lambda_1^2 N^{2-mD_1} + c_2\, \lambda_2'\mathscr{L}(N)\, \lambda_2\, N^{2-mD_1-D_2}$$

provided we assume

(7) $$0 < D_1 < 1/m, \quad mD_1 + D_2 < 1, \quad 0 < D_2 < 1.$$

Here c_1 and c_2 are constants depending on m, D_1 and D_2. Thus a choice of $\mathbf{B}_N$ here is

(8) $$\mathbf{B}_N = \begin{bmatrix} N^{H_1} & \mathbf{0}_{1\times p} \\ \mathbf{0}_{p\times 1} & N^{H_2}\mathbf{I}_{p\times p} \end{bmatrix} L(N) = \begin{bmatrix} b_{N1} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{N2} \end{bmatrix}, \text{ say}$$

with H_1 as in (5) and $2H_2 = 2 - mD_1 - D_2$.

With this $\mathbf{B}_N$, one can again verify that (6) above holds in this case. Note that (7) implies

(9) $$1/2 < H_1 < 1, \quad 1/2 < H_2 < 1.$$

Next, in view of (1.3), (1.5) and (3), $\forall\, \boldsymbol{\lambda} \in \mathscr{R}^{p+1}$,

$$E[\boldsymbol{\lambda}' \mathbf{B}_N^{-1} \sum_i \mathbf{X}_i \{\psi_1(\eta_i) - \frac{J_m}{m!} H_m(\eta_i)\}]^2 = E\{\sum_i \boldsymbol{\lambda}' \mathbf{B}_N^{-1} \mathbf{X}_i \sum_{q=m+1}^{\infty} \frac{J_q}{q!} H_q(\eta_i)]^2$$

$$= \sum_{q=m+1} \frac{J_q^2}{q!} \sum_i \sum_j E\, \boldsymbol{\lambda}' \mathbf{B}_N^{-1} \mathbf{X}_i \mathbf{X}_j' \mathbf{B}_N^{-1} \boldsymbol{\lambda} \cdot \rho^q(i-j)$$

$$\leq \sum_{q=m}^{\infty} \frac{J_q^2}{q!} \sum_i \sum_j \boldsymbol{\lambda}' \mathbf{B}_N^{-1} \cdot \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Gamma}(i-j) \end{bmatrix} \cdot \mathbf{B}_N^{-1} \boldsymbol{\lambda}\, |\rho^{m+1}(i-j)|.$$

(10) $$\longrightarrow 0,$$

by arguing as in Taqqu (1975, p. 294), under both (1.6a) of (1.6b), using $\mathbf{B}_N$ as in (5) or in (8), as the case may be.

Combining (5), (6), (8), and (10), one sees that under either (1.6a) or (1.6b) (with $\mathbf{B}_N$ as in (5) or as in (8), respectively) one has, by (3),

$$\mathrm{Var}(\boldsymbol{\lambda}' \mathbf{B}_N^{-1} \sum_i \mathbf{X}_i \psi_1(\eta_i))$$

$$= \mathrm{Var}[\sum_i (\boldsymbol{\lambda}' \mathbf{B}_N^{-1} \mathbf{X}_i)\{\psi_1(\eta_i) - \frac{J_m}{m!} H_m(\eta_i)\}] + \mathrm{Var}[\frac{J_m}{m!} \sum_i (\boldsymbol{\lambda}' \mathbf{B}_N^{-1} \mathbf{X}_i) H_m(\eta_i)]$$

$$= o(1) + O(1) = O(1).$$

This then determines $\mathbf{B}_N$ and verifies the assumption (b)(ii) of the Theorem 1 above when $\{\xi_i\}$, $\{\eta_i\}$, $\{\epsilon_i\}$ are as in (1.1), (1.4), (1.5), (1.6a) or (1.6b). Now, if $\mathbf{B}_N$ is given by (5), then

(11) $$A_N = \begin{bmatrix} N^{1-H} L(N) & \mathbf{0} \\ \mathbf{0} & N^{\frac{1}{2}} \mathbf{I}_{p\times p} \end{bmatrix} = \begin{bmatrix} a_{N1} & \mathbf{0} \\ \mathbf{0} & A_{N2} \end{bmatrix}, \text{ say,}$$

will satisfy (b)(i). If $\mathbf{B}_N$ is given by (8), then

(12) $$A_N = \begin{bmatrix} N^{1-H_1} & \mathbf{0} \\ \mathbf{0} & N^{1-H_2} \mathbf{I}_{p\times p} \end{bmatrix} L(N) = \begin{bmatrix} a_{N1} & \mathbf{0} \\ \mathbf{0} & A_{N2} \end{bmatrix}, \text{ say,}$$

will satisfy (b)(i), with H_1 and H_2 as in (5) and (8) satisfying (9).

The above discussion is now summarized as

***Theorem* 2.** *Let* $\{Y_i\}$, $\{\xi_i\}$, $\{\eta_i\}$, $\boldsymbol{\beta}$, ψ *satisfy* (1.1), (1.3), (1.4), (1.5), *and* (1.6a) or (1.6b). *In addition assume that* ψ *is nondecreasing satisfying* (a) *of Theorem* 1 *with* $0 < E\psi'(\epsilon)$. *Then, with* $\hat{\boldsymbol{\beta}}_N$ *defined as a solution of* (1.2),

$$A_N(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta}) = [\mathbf{R}\, E\psi'(\epsilon)]^{-1} \mathbf{B}_N^{-1} \sum_i \mathbf{X}_i H_m(\eta_i) \frac{J_m}{m!} + o_p(1).$$

where $\mathbf{B}_N$, $\mathbf{A}_N$ *are as in* (5), (11), ((8), (12)) *in the case of* (1.6a) ((1.6b)).

***Remark* 4. *Hermite rank* m *of* ψ_1.** Often the function ψ is chosen to be skew symmetric, viz, $\psi(-x) \equiv \psi(x)$. Thus if G is also skew symmetric then $\psi_1(-x) \equiv \psi(G(-x)) \equiv \psi(-G(x)) \equiv -\psi(G(x)) \equiv -\psi_1(x)$. In such cases, using the fact that $H_q(-x) \equiv (-1)^q H_q(x)$ for all q, we have

$$J_q = E\psi_1(\eta)H_q(\eta) = [1+(-1)^{q+1}]\cdot E\{\psi_1(\eta)H_q(\eta)I(\eta > 0)\} \neq 0, \quad q = 1.$$

Therefore, $m = 1$, $J_1 = 2\,E\{\psi_1(\eta)\eta I(\eta>0)\}$, $H_1(\eta) = \eta$ and, from Theorem 2,

$$A_N(\hat{\beta}_N-\beta) = [R\,E\psi'(\epsilon_1)]^{-1}\cdot B_N^{-1}\sum_i X_i\eta_i \cdot J_1+o_p(1). \tag{13}$$

Now let β_N be the least squares estimator of β in (1.1). Then carrying out an analysis like the above one can derive the following:

If $EG(\eta) = 0,\ 0 < EG^2(\eta) < \infty$ *and* G *is skew symmetric, then*

$$A_N(\beta_N - \beta) = R^{-1}\cdot B_N^{-1}\sum_i X_i\eta_i \cdot \alpha_1 + o_p(1),$$

where $\alpha_1 = EG(\eta)\eta$ where A_N and B_N are the same as in (13).

The r.v. $\Sigma X_i\eta_i$ is precisely the leading term in the least squares estimator of the regression parameter with the errors $\{\eta_i\}$ and the design vectors $\{X_i\}$. Thus it follows that the above class of M – estimators corresponding to the skew symmetric scores and symmetric errors are asymptotically like the least squares estimators regardless of whether the errors are Gaussian or not.

Now, suppose that there is no intercept parameter in (1.1). Then the result like (13), with X_i replaced by ξ_i, A_N, B_N replaced by A_{N2}, B_{N2} of (5) and/or (8) remains valid. Of course now $\hat{\beta}_N$ is $p\times 1$ as is β. Note that in the case of (1.6a),

$$\lambda' B_{N2}^{-1}\sum_i \xi_i\eta_i = N^{-\frac{1}{2}}\sum_i(\lambda'\xi_i)\eta_i \Rightarrow N_p(0, \lambda'\Gamma\lambda), \quad \forall\,\lambda \in \mathscr{R}^p,$$

where $\Gamma = \Gamma(0) = E\xi_1\xi_1'$.

But in the case of (1.6b) the limiting distribution is different. To determine this limiting distribution, we use Theorem 6.1 of Fox and Taqqu (1987). Observe that if $\{\xi_i\}$ are long range dependent and **Gaussian** then so are the r.v.'s $\{\lambda' \xi_i\}$ for every $\lambda \in \mathcal{R}^p$ with the same exponent D_2 as in (1.6b). Now, take X_i and Y_i in Theorem 6.1 of Fox and Taqqu to be $\lambda' \xi_i$ and η_i, respectively. One then sees that (1.4), (1.5) and (1.6b) together with the Gaussianness assumptions imply all the conditions of that Theorem for every λ. Hence,

$$\lambda' B_{N2}^{-1} \sum_i \xi_i \eta_i = N^{-H_2} L(N) \sum_i (\lambda' \xi_i)\eta_i => Z(1)\cdot(\lambda' \Gamma \lambda)^{\frac{1}{2}}$$

with Z(1) obtained from (6.1) of Fox and Taqqu after t is set equal to 1 in there. □

3. ***The Case of Non – Random Designs.*** In order to seperate this case from that of the random designs, we shall now denote an $N\times p$ design matrix of known constants by $\mathbf{C}$ and its ith row by $\mathbf{c}_{Ni}'$, $1 \le i \le N$. Consider the linear regression model where one observes $\{Y_{Ni}\}$ satisfying

$$(1) \qquad Y_{Ni} = \mathbf{c}_{Ni}'\boldsymbol{\beta} + \epsilon_i, \qquad 1 \le i \le N, \quad \boldsymbol{\beta} \in \mathcal{R}^p,$$

with $\{\epsilon_i\}$ as in (1.1).

Throughout we shall assume that

$$(L1) \qquad (\mathbf{C}'\mathbf{C})^{-1} \text{ exists for all } N \ge p.$$

The class of M – estimators $\hat{\boldsymbol{\beta}}_N$ is defined as a solution $\mathbf{t}$ of

$$(2) \qquad \mathbf{T}(\mathbf{t}) := \sum_i \mathbf{c}_{Ni}\psi(Y_{Ni} - \mathbf{c}_{Ni}'\mathbf{t}) = \mathbf{0},$$

where ψ is assumed to satisfy (1.3). Again, our objective here is to investigate the large sample behavior of these estimators when $\{\eta_i\}$ satisfy (1.4). Of course conceptually the discussion that follows is similar to that in Section 2 above except for the difficulties created by the nonstationarity that is introduced in the problem by $\{c_{Nj}\}$. We begin by giving

Theorem 1. *Let* $\epsilon_1, \epsilon_2, \ldots$ *be a strictly stationary sequence of r.v's and* C *be as above satisfying* (L1) *and assume* (1) *above holds. In addition, assume that the following hold:*

(L2) *The score function* ψ *is absolutely continuous with its almost everywhere derivative* ψ' *satisfying* $E|\psi'(\epsilon)| < \infty$ *and such that the function*

$z \longrightarrow E|\psi'(\epsilon - z) - \psi'(\epsilon)|$ *is continuous at zero.*

(L3) *There exists sequences* $\{A_N\}$ *and* $\{B_N\}$ *of* $p \times p$ *matrices such that they are positive definite for sufficiently large* N *and satisfy*

(i) $\|A_N^{-1}\| \longrightarrow 0$, $\|B_N^{-1}\| \longrightarrow 0$; (ii) $B_N^{-1} C' C A_N^{-1} = I_{p\times p}$

(iii) $\max_i \|A_N^{-1} c_{Ni}\| \longrightarrow 0$, (iv) $\|B_N^{-1} T(\beta)\| = O_p(1)$.

Then, for every $0 < b < \infty$,

$$(3)\qquad E \sup_{\|\Delta\| \le b} \|B_N^{-1} [T(\beta + A_N^{-1}\Delta) - T(\beta)] + B_N^{-1} \sum_i c_{Ni} c_{Ni}' \psi'(\epsilon_i) A_N^{-1} \Delta\| = o(1).$$

If, in addition, $\epsilon_i \equiv G(\eta_i)$, *with* $\{\eta_i\}$ *satifying* (1.4),

(L4) ψ *is nondecreasing,* $0 < E\psi'(\epsilon)$, $E(\psi'(\epsilon))^2 < \infty$, *and*

$$(L5)\qquad N^{1-(D/2)} \max_{1 \le i \le N} \|B_N^{-1} c_{Ni}\| \cdot \|A_N^{-1} c_{Ni}\| \longrightarrow 0, \quad \textit{with } D = D_1 \text{ of (1.4)}$$

then

$$(4)\qquad B_N^{-1} \cdot \sum_i c_{Ni} c_{Ni}' \, \psi'(\epsilon_i) \cdot A_N^{-1} = I_{p\times p} \cdot E\psi'(\epsilon) + o_p(1),$$

and

$$(5)\qquad A_N(\hat{\beta}_N - \beta) = [E\psi'(\epsilon)]^{-1} \cdot B_N^{-1} T(\beta) + o_p(1).$$

Remark 1. Some comments about the assumptions are in order. The assumptions (L2) and (L3) are similar to the assumptions (a) and (b) of Theorem 2.1 above. Recall that in the linear regression model with independent or weakly dependent errors and with the design matrix C, the magnitude of $\mathbf{T}(\beta)$ is of the order $\delta_N := (C'C)^{\frac{1}{2}}$. However in the current situation, where $\{\epsilon_i\}$ are functions of long range dependent r.v.'s, we can not expect this magnitude. But we must still have (L3)(ii) in order to stabilize the LHS(4).

In the case of random and stationary design variables, as in Section 2 above, an analogue of (4) is given by the Ergodic Theorem which does not require the second moment of the summands. But in the present situation, the LHS of (4) is neither stationary nor independent. The assumptions (L4), (L5) and (L3)(ii) together with the Gaussianness of $\{\eta_i\}$ is used to conclude (4) below.

Proof. To simplify writing, let $\mathbf{a}_i := A_N^{-1}c_{Ni}$, $\mathbf{b}_i := B_N^{-1}c_{Ni}$, $1 \le i \le N$. Now, by the absolute continuity of ψ, the Fubini Theorem and the Cauchy–Schwarz inequality, the

$$\mathrm{LHS}(3) \le 2b \sum_i \|\mathbf{b}_i\| \, \|\mathbf{a}_i\| \, \{2\|\mathbf{a}_i\|\}^{-1} \int_{-\|\mathbf{a}_i\|}^{\|\mathbf{a}_i\|} E|\psi'(\epsilon - zb) - \psi'(\epsilon)|\,dz$$

$$\le 2b\Big(\sum_i \|\mathbf{b}_i\|^2 \sum_i \|\mathbf{a}_i\|^2\Big)^{\frac{1}{2}} \times \max_i\Big[(2\|\mathbf{a}_i\|)^{-1} \int_{-\|\mathbf{a}_i\|}^{\|\mathbf{a}_i\|} E|\psi'(\epsilon - zb) - \psi'(\epsilon)|\,dz\Big] \longrightarrow 0,$$

by (L2), (L3)((i) – (iii)). Note that by (L3)(ii),

$$\sum_i \|\mathbf{b}_i\|^2 \sum_i \|\mathbf{a}_i\|^2 = \mathrm{tr.}B_N^{-1}C'CB_N^{-1} \cdot A_N^{-1}C'CA_N^{-1} = p = O(1)$$

where $\mathrm{tr.}A := \mathrm{trace}\ A$ for any matrix A.

Next, let $\psi_2(\eta) := \psi'(\epsilon) = \psi'(G(\eta))$ and $\alpha_q := E\psi_2(\eta)H_q(\eta)$. In view of (L4), the Hermite expansion of $\psi_2(\eta_i) - E\psi_2(\eta)$ is $\sum_{q=1}^{\infty} \frac{\alpha_q}{q!} H_q(\eta_i)$. Also note that the LHS(4) above is now $\sum_i \mathbf{b}_i\mathbf{a}_i' \, \psi_2(\eta_i)$. Hence $\forall\ \boldsymbol{\lambda} \in \mathcal{R}^p$,

$$E\|\boldsymbol{\lambda}' \sum_i \mathbf{b}_i\mathbf{a}_i'[\psi_2(\eta_i) - E\psi_2(\eta)]\|^2 = E\|\boldsymbol{\lambda}' \sum_i \mathbf{b}_i\mathbf{a}_i' \sum_{q=1}^{\infty} \frac{\alpha_q}{q!} H_q(\eta_i)\|^2$$

$$= \sum_{q=1}^{\infty} \frac{\alpha_q^2}{q!} \sum_i \sum_j \lambda' b_i a_i' a_j b_j' \lambda \cdot \rho^q(i-j) \qquad \text{by (2.3)},$$

$$\leq \operatorname{Var}(\psi_2(\eta)) \cdot \|\lambda\|^2 \max_i \|b_i a_i'\|^2 \sum_i \sum_j |\rho(i-j)|.$$

(6) $$= \max_i \{\|B_{Ni}^{-1} c_{Ni}\| \cdot \|A_N^{-1} c_{Ni}\|\}^2 \cdot O(N^{2-D}),$$

because $\sum_i \sum_j |\rho(i-j)| = O(N^{2-D})$ and because

$$\|b_i a_i'\|^2 \equiv \operatorname{tr}\cdot(b_i a_i' a_i b_i') = \operatorname{tr}.(b_i' b_i)\cdot(a_i' a_i)$$

$$\equiv \|b_i\|^2 \|a_i\|^2 \equiv \|B_N^{-1} c_{Ni}\|^2 \|A_N^{-1} c_{Ni}\|^2.$$

Therefore (4) follows from the assumption (L4) and (6). The result (5) follows as in Huber (op cit.). □

Our next objective is to determine B_N, using (L3)(iv). Again, to simplify exposition we shall write $\{c_i\}$ for $\{c_{Ni}\}$. Proceeding as in Section 2, we observe that $\forall\ \lambda \in \mathscr{R}^p$,

$$E\{\lambda' \sum_i c_i \psi_1(\eta_i)\}^2 = E\{\lambda' \sum_i c_i H_m(\eta_i) \cdot \frac{J_m}{m!}\}^2 + E\{\lambda' \sum_i c_i[\psi_1(\eta_i) - H_m(\eta_i) \cdot \frac{J_m}{m!}]\}^2$$

$$= E\{\lambda' \sum_i c_i H_m(\eta_i) \cdot \frac{J_m}{m!}\}^2 + E[\lambda' \sum_i c_i \sum_{q \geq m+1} \frac{J_q}{q!} H_q(\eta_i)]^2$$

$$= \lambda' \sum_i \sum_j c_i c_j' \rho^m(i-j)\, \lambda \cdot \frac{J_m}{m!} + \sum_{q \geq m+1} \frac{J_q}{q!} \sum_i \sum_j \lambda' c_i c_j' \lambda \cdot \rho^q(i-j)$$

(7) $$= \frac{J_m}{m!} \lambda' K_{N1} \lambda + \lambda' K_{N2} \lambda$$

where

$$K_{N1} := \sum_i \sum_j c_i c_j' \rho^m(i-j), \quad K_{N2} := \sum_{q \geq m+1} \frac{J_q^2}{q!} \sum_i \sum_j c_i c_j' \rho^q(i-j).$$

At this point one is clearly persuaded to choose $B_N \approx K_{N1}^{\frac{1}{2}}$ and then try to show that $\|K_{N1}^{-\frac{1}{2}} K_{N2} K_{N1}^{-\frac{1}{2}}\| \to 0$ so that we would have (L3)(iv) satisfied. Such a process, though feasible, appears to be quite involved for general $\{c_i\}$. However, if we make some further assumptions on the design variable then this process is less involved and more transparant.

Accordingly, let $\varphi^t := (\varphi_1,\ldots,\varphi_p)$ be a vector of measurable functions on $[0,1]$ to $\mathscr{R}$ satisfying the following conditions:

(a1) With $D = D_1$ and L as in (1.4), m as the Hermite rank of $\psi_1(\eta)$,

(i) $\int_0^1 \int_0^{1-u} |\varphi_\ell(u)\, \varphi_k(u+v) v^{-mD} L(v)|\, dv\, du < \infty, \qquad D < 1/m,$

(ii) $\int_0^1 |\varphi_\ell(u)\, \varphi_k(u)|\, du < \infty, \qquad \ell,k = 1,2,\ldots,p.$

(a2) (i) $N^{-D/4} \max_{1 \le i \le N} \|\varphi(i/N)\| \to 0$; (ii) $N^{-1+mD} \max_{1 \le i \le N} \|\varphi(i/N)\|^2 \to 0$.

(a3) The matrix $\mathscr{G}^{-1}$ exists, where

$$\mathscr{G} = ((g_{\ell k})),\ g_{\ell k} = \int_0^1 \int_0^1 \varphi_\ell(u)\varphi_k(v)|v-u|^{-mD} L(|v-u|) du\, dv,\ \ \ell, k = 1,\ldots,p.$$

Given such a collection of φ's, choose

$$c_i := \varphi(i/N),\ 1 \le i \le N. \tag{8}$$

Now observe that

$$N^{-1} C'C = N^{-1} \sum_i c_i c_i' \approx \int_{1/N}^{1-1/N} \varphi(u)\varphi^t(u) du \to \int_0^1 \varphi(u)\varphi^t(u) du,$$

so that

$$N^{-2+mD}\, C'C \to 0, \qquad \text{because } -1 + mD < 0. \tag{9}$$

From (1.4), (8) and the slowly varying property of L it follows that

$$\lambda' K_{N1} \lambda = \sum_{\ell=1}^{p} \sum_{k=1}^{p} \lambda_\ell \lambda_k \sum_i \sum_j \varphi_\ell(i/N)\varphi_k(j/N)\, \rho^m(j-i)$$

$$= \lambda' C' C \lambda + 2 \sum_{\ell=1}^{p} \sum_{k=1}^{p} \lambda_\ell \lambda_k \sum_{i<j} \sum \varphi_\ell(i/N)\varphi_k(j/N)\, \rho^m(j-i)$$

$$\approx N^{2-mD}\, 2 \sum_{\ell=1}^{p} \sum_{k=1}^{p} \lambda_\ell \lambda_k \int_0^1 \int_0^{1-u} \varphi_\ell(u)\varphi_k(u+v)v^{-mD}L(v)dudv$$

$$= N^{2-mD}\, \lambda' \mathcal{G} \lambda.$$

Now let

$$\text{(10)} \qquad B_N := N^H \mathcal{G}^{1/2}, \quad H = 1-(mD/2), \quad D = D_1 \text{ of (1.4)}.$$

Our next objective is to show that the second term in the RHS(7) is $O(N^{2H})$. To that effect, note that $q \geq m+1$, $|\rho(k)| \leq 1$, $\forall k \geq 1$, imply that

$$\text{(11)} \qquad |\lambda' \sum_{i<j} \sum c_i c_j' \rho^q(j-i)\, \lambda| \leq \sum_{\ell=1}^{p} \sum_{k=1}^{p} |\lambda_\ell \lambda_k| \sum_{i<j} \sum |\varphi_\ell(i/N)\varphi_k(j/N)\, \rho^{m+1}(j-i)|.$$

Now, since $|\rho(k)| \to 0$ as $k \to \infty$, $\forall\, \epsilon > 0\ \exists\, N_\epsilon$ such that $|\rho(k)| \leq \epsilon\ \forall\, k > N_\epsilon$. Hence, $\forall\, N > N_\epsilon$,

$$\sum_{i<j} \sum |\varphi_\ell(i/N)\varphi_k(j/N)\, \rho^{m+1}(j-i)| \leq \sum_{|j-i| \leq N_\epsilon} \sum |\varphi_\ell(i/N)\varphi_k(j/N)|$$

$$+ \epsilon \cdot \sum_{i<j;(j-i)>N_\epsilon} \sum |\varphi_\ell(i/N)\varphi_k(j/N)\, \rho^m(j-i)|$$

$$= T_{N1} + \epsilon \cdot T_{N2}, \qquad \text{say.}$$

But, $\forall\, \ell, k = 1,\ldots,p$,

$$T_{N1} \leq N_\epsilon \cdot N \cdot \max_i \|\varphi(i/N)\|^2 = o(N^{2-mD}), \quad \text{by (a2)(ii)},$$

$$T_{N2} \leq \sum_{i<j} \sum |\varphi_\ell(i/N)\varphi_k(j/N)|\ |\rho^m(j-i)|$$

$$\approx N^{2-mD} \int_0^1 \int_0^{1-u} |\varphi_\ell(u)\varphi_k(v+u)v^{-mD}L(v)|du\, dv = O(N^{2-mD}), \text{ by (a1)(i).}$$

Hence, $\forall\ \epsilon > 0,\ \exists\ N_\epsilon$ such that

$$\text{LHS}(11) \leq o(N^{2-mD}) + \epsilon \cdot O(N^{2-mD}), \quad \forall\ N > N_\epsilon. \tag{12}$$

From (9), (12) and the definition of K_{N2} it follows that $\forall\ N > N_\epsilon$,

$$\begin{aligned} N^{-2+mD}\ |\lambda' K_{N2}\ \lambda| &\leq \text{Var}\ \psi_1(\eta) \cdot \{|\lambda' C' C\ \lambda| + \text{LHS}(11)\} \\ &\leq o(1) + \epsilon \cdot O(1) \longrightarrow 0, \quad \text{by now letting } \epsilon \longrightarrow 0. \end{aligned}$$

It thus follows that (L3)(iv) holds with B_N given by (10). From (L3)(ii) we get

$$A_N = B_N^{-1} \cdot C' C \approx N^{1-H} \mathscr{G}^{-1/2} \int_0^1 \varphi\, \varphi^t. \tag{13}$$

Note that $m \geq 1$ =>

$$\max_i \|A_N^{-1} c_i\| \approx \max_i N^{-mD/2} \|\varphi(i/N)\| \leq N^{-D/4} \max_i \|\varphi(i/N)\|$$

and

$$N^{1-D/2} \max_i [\|A_N^{-1} c_{Ni}\| \cdot \|B_N^{-1} c_{Ni}\|] \approx [N^{-D/4} \max_i \|\varphi(i/N)\|]^2$$

so that (a2) implies (L3)(ii) and (L5). This shows that all the assumptions of Theorem 1 are satisfied. We now summarize the above discussion as

***Theorem* 2.** *Suppose that the linear regression model* (1), *with errors as in* (1.1) *and* (1.4), *holds. About the design variables* $\{c_{Ni}\}$ *and the score function* ψ *assume that* (8), (a1)–(a3), (1.3), (L2) *and* (L4) *hold. Then* M – estimators $\{\hat{\beta}_N\}$ *defined as solutions of* (2) *satisfy*

$$N^{1-H}(\hat{\beta}_N - \beta) = \{m! \cdot \int_0^1 \varphi\varphi^t \cdot E\psi'(\epsilon)\}^{-1} \cdot N^{-H} \sum_i \varphi(i/N) H_m(\eta_i) \cdot J_m + o_p(1), \tag{14}$$

where $H = (1-mD/2)$, $D = D_1$ *of* (1.4).

***Remark* 2.** Observe that if the design generating functions are bounded then $0<D<1/m$ guarantees the satisfaction of (a1) and (a2). In particular if $\varphi_\ell(u) = u^\ell$, $\ell = 1,\dots,p$, then (a1) – (a3) are all satisfied. That is, all of these conditions are satisfied in the case of the p^{th} order polynomials.

An example of an unbounded design is obtained by taking p=1, $\varphi_1(u) = u^{-r}$, r>0. Then (a1)–(a3) are satisfied as long as $r < (1-mD)/2$.

***Remark* 3.** An analogue of Remark 2.4 applies here also with obvious modifications. Consequently, for skew symmetric ψ and symmetric errors the asymptotic distribution of $N^{1-H}(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta})$ is p – variate Normal with mean vector **0** and the covariance matrix $[\{\int_0^1 \varphi\varphi^t\}^{-1} \boldsymbol{\mathcal{G}} \{\int_0^1 \varphi\varphi^t\}^{-1}] \cdot \{E\psi'(\epsilon)\}^{-2} J_1^2$.

Acknowledgement. Author would like thank Murad Taqqu for providing a generous supply of preprints and reprints of his and his co–authors work on long range dependent processes. □□

REFERENCES

Beran, J. and Kunsch, H.–R. (1985). Location estimators for processes with long range dependence. ETH Zurich, Res. Rep. #40.

Dehling, H. and Taqqu, M.S. (1989). The empirical process of some long range dependent sequences with an Application to U–statistics. *Ann. Statist.*, **17**, 1767 – 1783.

Feller, W. (1971). *An Introduction to Probability Theory and Its Applications.* J. Wiley & Son. New York, NY.

Fox, R. and Taqqu, M.S. (1987). Multiple stochastic integrals with dependent integrators. *J. Multi. Analys.* **21**, 105–127.

Granger, C.W.J. and Joyeux, R. (1980). An introduction to long–memory time series models and fractional differencing. *J. Time Ser. Anal.* **1**, 15–29.

Hosking, J.R.M. (1981). Fractional differencing. *Biometrika*, **68**, 165–176.

Huber, P.J. (1981). *Robust Statistics.* John Wiley & Sons, New York.

Koul,H.L. (1977). Behavior of robust estimators in regression model with dependent errors. *Ann. Statist.*, **5**, 681–699.

Mandelbrot, B.B. and Van Ness, J.W. (1968). Fractional Brownian motions, fractional noise and applications. *SIAM Rev.* **10**, 422–437.

Taqqu, M.S. (1975). Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahr. Geb.* **31**, 287–302.

Taqqu, M.S. (1979). Self–similar processes and related ultraviolet and infrared catastrophes. *Colloq. Math. Soc. Janos Bolyai: Random Fields*, **27**, 1057–1096.

Yajima, Y. (1988). On estimation of a regression model with long–memory stationary errors. *Ann. Statist.* **16**, 791–807.

Yajima, Y. (1985). On estimation of long–memory time series models. *Austral. J. Statist.* **27**, 303–320.

256. E. Mayer-Wolf, A central limit theorem in nonlinear filtering, Apr. 89.

257. C. Houdré, Factorization algorithms and non-stationary Wiener filtering, Apr. 89, Stochastics, to appear.

258. C. Houdré, Linear Fourier and stochastic analysis, Apr. 89.

259. G. Kallianpur, A line grid method in areal sampling and its connection with some early work of H. Robbins, Apr. 89. *Amer. J. Math. Manag. Sci.*, 1989, to appear.

260. G. Kallianpur, A.G. Miamee and H. Niemi, On the prediction theory of two-parameter stationary random fields, Apr. 89. *J. Multivariate Anal.*, to appear.

261. I. Herbst and L. Pitt, Diffusion equation techniques in stochastic monotonicity and positive correlations, Apr. 89.

262. R. Selukar, On estimation of Hilbert space valued parameters, Apr. 89. (*Dissertation*)

263. E. Mayer-Wolf, The noncontinuity of the inverse Radon transform with an application to probability laws, Apr. 89.

264. D. Monrad and W. Philipp, Approximation theorems for weakly dependent random vectors and Hilbert space valued martingales, Apr. 89.

265. K. Benhenni and S. Cambanis, Sampling designs for estimating integrals of stochastic processes, Apr. 89.

266. S. Evans, Association and random measures, May 89.

267. H. Hurd, Correlation theory of almost periodically correlated processes, June 89.

268. O. Kallenberg, Random time change and an integral representation for marked stopping times, June 89.

269. O. Kallenberg, Some uses of point processes in multiple stochastic integration, Aug. 89.

270. W. Wu and S. Cambanis, Conditional variance of symmetric stable variables, Sept. 89.

271. J. Mijnheer, U-statistics and double stable integrals, Sept. 89.

272. O. Kallenberg, On an independence criterion for multiple Wiener integrals, Sept. 89.

273. G. Kallianpur, Infinite dimensional stochastic differential equations with applications, Sept. 89.

274. G.W. Johnson and G. Kallianpur, Homogeneous chaos, p-forms, scaling and the Feynman integral, Sept. 89.

275. T. Hida, A white noise theory of infinite dimensional calculus, Oct. 89.

276. K. Benhenni, Sample designs for estimating integrals of stochastic processes, Oct. 89. (*Dissertation*)

277. I. Rychlik, The two-barrier problem for continuously differentiable processes, Oct. 89.

278. G. Kallianpur and R. Selukar, Estimation of Hilbert space valued parameters by the method of sieves, Oct. 89.

279. G. Kallianpur and R. Selukar, Parameter estimation in linear filtering, Oct. 89.

280. P. Bloomfield and H. Hurd, Periodic correlation in stratospheric ozone time series, Oct. 89.

281. J.M. Anderson, J. Horowitz and L.D. Pitt, On the existence of local times: a geometric study, Jan. 90.

282. G. Lindgren and I. Rychlik, Slepian models and regression approximations in crossing and extreme value theory, Jan. 90.

283. H.L. Koul, M-estimators in linear models with long range dependent errors, Feb. 90.