THE INSTITUTE OF STATISTICS

THE CONSOLIDATED UNIVERSITY OF NORTH CAROLINA

On the Asymptotic Distribution of Multivariate M-Estimates

Raymond J. Carroll

Institute of Statistics Mimeo Series No. 1134
September, 1977

DEPARTMENT OF STATISTICS
Chapel Hill, North Carolina

On the Asymptotic Distribution of Multivariate M-Estimates

by

Raymond J. Carroll*

University of North Carolina at Chapel Hill

ABSTRACT

The asymptotic distribution of multivariate M-estimates is studied. It is shown that, in general, consistency leads to asymptotic normality and a Law of the Iterated Logarithm. The results are used to compute via matrix derivatives the asymptotic distribution of a class of estimates due to Maronna.

AMS 1970 Subject Classification: Primary 62G35; Secondary 62E20.

Key Words and Phrases: Matrix derivatives, robust estimation, jackknife, multivariate analysis, M-estimators.

*This work was supported by the Air Force Office of Scientific Research under contract AFOSR-76-2796.

Introduction

This paper examines the asymptotic distribution of multivariate M-estimates, with emphasis on a recent proposal of Maronna (1976). In section 2, the necessary general theory is given. The methods of Carroll (1977) and Huber (1967) are used to show that, under some conditions, consistency of M-estimates implies asymptotic normality, a Law of the Iterated Logarithm (LIL), and an approximation of M-estimates by the sample mean of bounded random variables, with the error of approximation of almost sure order $O(n^{-1}\log_2 n)$.

Maronna (1976) computed for distributions of radial type (of dimension p) the asymptotic distribution of M-estimates of location and scatter (say T_n, V_n), defined as solutions of systems of equations of the form

$$n^{-1}\sum_{i=1}^{n}\rho_1(\{(X_i-t)'V^{-1}(X_i-t)\}^{1/2})(X_i-t) = 0 \tag{1.1}$$

$$n^{-1}\sum_{i=1}^{n}(X_i-t)\rho_2((X_i-t)'V^{-1}(X_i-t))(X_i-t)' = V \ , \tag{1.2}$$

where ρ_1 and ρ_2 are smooth functions. Under his assumptions, T_n and V_n are asymptotically independently and normally distributed when properly normed, and T_n has a particularly simple covariance matrix. In section 3 we investigate the relationship of T_n and V_n in a formal manner under weaker distributional assumptions, although under stronger restrictions for ρ_1 and ρ_2. Since Maronna has shown the consistency of his estimates under conditions weaker than radial symmetry, the results of our section 2 can be used to investigate the exact nature of the asymptotic distributions of T_n and V_n when the underlying distribution is not of radial type. This investigation is facilitated by the use of Taylor expansions with matrix derivatives (MacRae (1974)). The power and case of matrix derivatives have greatly simplified the computations.

Qualitatively, our results (particularly (3.4)) exhibit the phenomenon that the asymptotic covariance matrix Σ of T_n is a complicated function depending on V_n, except in cases of symmetry. This means that direct estimation of Σ from an asymptotic variance formula (as is

done in the symmetric case), while possible, is not appealing. Fortunately, T_n is a smooth functional of ρ_1 and ρ_2, so that jackknifing would be an appropriate method of estimation; however, T_n is a nonlinear function of $X_1,X_2,\ldots,X_n$, so that the jackknife will be rather burdensome from a computational standpoint.

Representations

In this section we show that consistency of M-estimates leads, under general smoothness conditions, to a much stronger characterization. We adopt the notation of Huber (1967), so that Θ will be a connected subset of p-dimensional Euclidean space $\mathbb{R}^p$, (X,Ω,P) is the probability space, $X = \mathbb{R}^q$ for some q, and $\psi(x,\theta)$ is a function mapping $X\times\Theta$ into $\mathbb{R}^p$. ψ will be given the measurability and separability by Huber's (B-1). We define a sequence of statistics T_n in such a way that

$$n^{-1}\textstyle\sum_1^n\psi(X_i,T_n) = 0 \ , \qquad (2.1)$$

where $X_1,X_2,\ldots$ is a sequence of i.i.d. random vectors with common probability distribution P. We wish to investigate the properties of any consistent sequence of solutions to (2.1), so we assume the existence of θ_0 with $T_n \to \theta_0$ almost surely, where θ_0 is an isolated (not necessarily unique) zero of

$$E\psi(X,\theta) = 0 \ . \qquad (2.2)$$

Maronna (1976) has a multivariate example which under certain conditions has a unique zero to (2.2), while, in another related context, Collins (1976) constructs an algorithm for finding T_n even though the zero to (2.2) is not unique. One step of a Newton-Raphson iteration is also a possibility (Bickel (1974)).

We assume in this section that for each θ, $\psi(x,\theta)$ has second partial derivatives in θ except possibly for at most k points $X = a_1(\theta),\ldots,a_k(\theta)$. Let $B(\theta,\varepsilon)$ be the union of k open squares of length ε which have centers at $a_1(\theta),\ldots,a_k(\theta)$ and let $B(\varepsilon)$ be the closure of the set

$$\cup\{B(\theta,\varepsilon):\ ||\theta-\theta_0|| \le \varepsilon\}\ ,$$

where $||\theta||$ is the maximum absolute component of θ. We assume $B(\varepsilon)$ is compact in $\mathbb{R}^q = X$.

The following result generalizes one given by Carroll (1977), who studied M-estimates in the one-dimensional location case with preliminary estimate of scale. The proofs are similar, so that only the important details are given here.

<u>Theorem 1</u>. For some $\varepsilon > 0$, suppose that for $||\theta-\theta_0|| \le \varepsilon$, the following hold:

(2.3) There exists a constant M such that if $\varepsilon' \le \varepsilon$, then

$$\Pr\{X_1 \epsilon B(\theta,\varepsilon')\} \le M\varepsilon'\ .$$

(2.4) The matrix $A = E\partial\psi(X,\theta_0)/\partial\theta'$ exists and is of full rank.

(2.5) On $B(\varepsilon)$, if $||\theta-\theta_0|| \le \varepsilon$, then $||\psi(x,\theta) - \psi(x,\theta_0)|| \le M||\theta-\theta_0||$. Further, $\partial\psi(x,\theta)/\partial\theta'$ is bounded on $B(\varepsilon) - \{a_1(\theta_0),\ldots,a_k(\theta_0)\}$.

(2.6) ψ has second partial derivatives in θ with

$$\sup\{||\ \frac{\partial^2\psi(x,\theta)}{\partial\theta\partial\theta'}\ ||:\ ||\theta-\theta_0|| \le \varepsilon,\ x \ne a_1(\theta),\ldots,a_k(\theta)\} < \infty\ .$$

Then, as $n \to \infty$, there is a constant C such that

(2.7) $$T_n-\theta_0 = A^{-1}(n^{-1}\sum_{i=1}^{n}\psi(X_i,\theta_0)) + G_n,\ \text{where}\ \lim_n\sup n(\log_2 n)^{-1}||G_n|| \le C\quad \text{(a.s)}.$$

The heart of the proof is contained in the following lemma.

<u>Lemma 1</u>. Under the conditions of Theorem 1, there exists a positive number C such that for any sequence $\varepsilon_1,\varepsilon_2,\ldots$ in $[0,\varepsilon]$ the following holds under P:

almost surely as $n \to \infty$ for $||\theta-\theta_0|| \le \varepsilon_n$,

$$(2.8) \qquad ||A^{-1}n^{-1}\sum_{i=1}^{n}\psi(X_i,\theta) - A^{-1}n^{-1}\sum_{i=1}^{n}\psi(X_i,\theta_0) - (\theta-\theta_0)||$$

$$\le C\{||\theta-\theta_0|| + n^{-1/2}(\log_2 n)^{1/2}\}^2 + C||\theta-\theta_0||\max\{\varepsilon_n,(\log n)/n\} ,$$

$$(2.9) \qquad A^{-1}n^{-1}\sum_{i=1}^{n}\psi(X_i,\theta) - (\theta-\theta_0)|| \le C\{||\theta-\theta_0||\varepsilon_n + n^{-1/2}(\log_2 n)^{1/2}\} .$$

Proof of Lemma 1. Suppose $X\in B(\varepsilon_n)$ but $X \ne a_j(\theta_0)$ for any $j=1,\ldots,k$. By (2.5),

$$(2.10) \qquad ||\psi(X,\theta) - \psi(X,\theta_0) - A(\theta-\theta_0)|| \le C_1||\theta-\theta_0||I_{B(\varepsilon_n)}(X) ,$$

where $I_B(\cdot)$ is the indicator function of the event B. Similarly, for X outside $B(\varepsilon_n)$, by a Taylor expansion and (2.6), the left hand side of (2.10) is bounded by $C_2||\theta-\theta_0||^2$. Hence

$$||\psi(X,\theta) - \psi(X,\theta_0) - A(\theta-\theta_0)||$$
$$\le C_3||\theta-\theta_0||\{||\theta-\theta_0|| + I_{B(\varepsilon_n)}(X)\} .$$

Now, since $B(\varepsilon)$ is compact, it has a finite subcover so that by (2.3), $\Pr\{X_1\in B(\varepsilon_n)\} \le M_1\varepsilon_n$ for some constant M_1. Then, one shows by Bernstein's inequality that as $n \to \infty$,

$$\limsup_n \frac{n^{-1}\sum_1^n I_{B(\varepsilon_n)}(X_i)}{\max(\varepsilon_n, n^{-1}\log n)} \le 4 \quad \text{(a.s.)} .$$

Also, by the Law of the Iterated Logarithm and the second part of (2.5), almost surely as $n \to \infty$

$$||\left(n^{-1}\sum_1^n \frac{\partial\psi(X_i,\theta_0)}{\partial\theta'} - A\right)(\theta-\theta_0)||$$
$$\le C_4||\theta-\theta_0||n^{-1/2}(\log_2 n)^{1/2} .$$

Thus, (2.8) follows. To prove (2.9), note that since ψ is bounded, by the Law of the Iterated Logarithm, almost surely as $n \to \infty$,

$$A^{-1}n^{-1}\sum_1^n\psi(X_i,\theta) - (\theta-\theta_0)|| \le ||A^{-1}n^{-1}\sum_1^n\psi(X_i,\theta_0)||$$

$$+ ||A^{-1}n^{-1}\sum_1^n\psi(X_i,\theta) - A^{-1}n^{-1}\sum_1^n\psi(X_i,\theta_0) - (\theta-\theta_0)||$$

$$\le C_5\{||\theta-\theta_0||\varepsilon_n + n^{-1/2}(\log_2 n)^{1/2}\} . \qquad \square$$

Proof of Theorem 1. Take $\varepsilon_n \equiv \varepsilon' = \min(\varepsilon,(2C)^{-1})$. We may assume $||T_n-\theta_0|| \le \varepsilon'$, so that consider points θ for which at least one component of $\theta-\theta_0$ falls in one of the intervals

$$[C_4n^{-1/2}(\log_2 n)^{1/2},\varepsilon'] \quad \text{or} \quad [-\varepsilon', -C_4n^{-1/2}(\log_2 n)^{1/2}] ,$$

for a constant C_4. If C_4 is sufficiently large,

$$C\{||\theta-\theta_0||\varepsilon_n + n^{-1/2}(\log_2 n)^{1/2}\} \le ||\theta-\theta_0||/2 + C_4n^{-1/2}(\log_2 n)^{1/2}(C/C_4)$$

$$\le ||\theta-\theta_0||\{1/2 + (C/C_4)\} < ||\theta-\theta_0|| .$$

Hence, for n sufficiently large, if one component of $\theta-\theta_0$ falls in either of the intervals (2.11), then

$$(2.12) \qquad A^{-1}n^{-1}\sum_1^n\psi(X_i,\theta) - (\theta-\theta_0)|| < ||\theta-\theta_0|| .$$

Consider T_n. If its largest element falls in one of the two intervals, then by (2.12) and the relation

$$n^{-1}\sum_1^n\psi(X_i,T_n) = 0$$

we obtain $||T_n-\theta_0|| < ||T_n-\theta_0||$, a contradiction. Thus, as $n \to \infty$, $||T_n-\theta_0|| \le C_4n^{-1/2}(\log_2 n)^{1/2}$ (a.s.). Now place T_n in the expansion (2.9) to complete the proof.

The i.i.d. assumption for $X_1,X_2,\ldots$ was used only to insure the LIL for certain sums and for use of Bernstein's inequality; thus the proof can be generalized to include cases of non-independence or non-stationarity. Note further that the results do not require complete differentiability of ψ , so that Theorem 1 includes the ρ_1 and ρ_2 in (1.1) and (1.2) which define Huber's Proposal 2 (Maronna (1976)).

Multivariate Location and Scatter

The estimates defined by (1.1) and (1.2) are studied in this section. As discussed in the introduction, the results of the preceeding section can typically be used to show that for some points $(T(P), V_0(P))$,

$$(3.1) \qquad \begin{aligned} T_n - T(P) &= O(n^{-1/2}(\log_2 n)^{1/2}) \quad \text{(a.s.)} , \\ T_n - V_0(P) &= O(n^{-1/2}(\log_2 n)^{1/2}) \quad \text{(a.s.)} , \end{aligned}$$

and this is assumed throughout. Also, Theorem 1 enables us to assume that ρ_1 and ρ_2 are sufficiently smooth as to allow Taylor expansions, which we will do throughout this section. We assume the matrices H_1, H_2, and H_3 defined below are invertible; this has been shown by Maronna when the distributions are of radial type.

The following notation is used. I_p denotes the $(p\times p)$ identity matrix $d^2(X,V) = X'V^{-1}X$, "$\otimes$" is the Kronecker product, while $A*B$ is the star produce (MacRae (1974)). If $W = (w_1,\ldots,w_p)$ is $(p\times p)$, then $W^* = (w_1',\ldots,w_p')$ is a $(p^2\times 1)$ matrix. If X is $(n\times m)$, the matrix derivative $\partial X/\partial X = E_{(n,m)}$ is an $(n^2\times m^2)$ matrix. The chain rule and product rule remain as defined by MacRae, while $\rho_3(x) = \rho_1'(x)/(2x)$. The following definitions presuppose that H_2^{-1} exists, with E_1 and E_{01} denoting expectations with respect to the measures of X_1 and (X_0,X_1) (independently distributed as P) respectively.

Definitions. Set

$$H_1 = E_1\{\rho_1(d(X_1,V_0))I_p + 2\rho_3(d(X_1,V_0))X_1X_1'V_0^{-1}\},$$

$$H_2 = \left\{I_{p^2} + E_1\{\rho_2'(d^2(X_1,V_0))(X_1X_1')^*(V_0^{-1}X_1X_1'V_0^{-1})^{*\prime}\}\right\},$$

$$(W(X))^{*\prime} = (V_0^{-1}XX'V_0^{-1})^{*\prime}H_2^{-1},$$

$$H_3 = H_1 - E_{01}\Big\{\rho_3(d(X_1,V_0))X_1X_0'\{\rho_2(d^2(X_0,V_0))(W(X_1) + W(X_1)') + \rho_2'(d^2(X_0,V_0))(2V_0^{-1})X_0X_0'W(X_1)\}\Big\}.$$

Theorem 2. Assume H_1^{-1}, H_2^{-1} exist and set $T(P) = 0$, $V_0(P) = V_0$.

Then, almost surely as $n\to\infty$,

$$(3.2)\quad H_1(T_n - T(P)) = H_1T_n = n^{-1}\sum_1^n X_i\rho_1(d(X_i,V_0)) - E_1\{\rho_3(d(X_1,V_0))X_1(V_0^{-1}X_1X_1'V_0^{-1})^{*\prime}\}(V_n-V_0)^* + O(n^{-1}(\log_2 n)),$$

$$(3.3)\quad H_2(V_n-V_0)^* = n^{-1}\sum_1^n\{X_i\rho_2(d^2(X_i,V_0))X_i' - V_0\}^* - E_1\{\rho_2(d^2(X_1,V_0))(T_nX_1' + X_1T_n')^*\} - 2E_1\{\rho_2'(d^2(X_1,V_0))(X_1X_1'V_0^{-1}T_nX_1')^*\} + O(n^{-1}\log_2 n),$$

$$(3.4)\quad H_3T_n = n^{-1}\sum_1^n X_i\rho_1(d(X_i,V_o)) - E_0\{\rho_3(d(X_0,V_0))X_0W(X_0)\}n^{-1}\sum_1^n\{X_i\rho_2(d^2(X_i,V_0))X_i' - V_0\}^* + O(n^{-1}\log_2 n).$$

In providing Theorem 2, we make use of the following identities (De Waal (1975)). In equations (3.5) to (3.11) below, X, W, and T will be $(p \times 1)$ column vectors, while V, W_1, and W_2 are $(p \times p)$ matrices.

(3.5) $$(ABC)' = B' * (C \otimes I_n) E_{(n,m)} (A \otimes I_m) \quad A(m \times p),\ B(p \times q),\ C(q \times m),$$

(3.6) $$W_1 * I_p = W',$$

(3.7) $$(X' \otimes I_p) + XE_{(1,p)})(I_p \otimes T) = TX' + XT',$$

(3.8) $$(X' \otimes I_p)(I_p \otimes T) = X' \otimes T = TX'$$

(3.9) $$W_1^{*\prime}(W_2 TX')^* = \mathrm{tr}(W_1 XT' W_2') = X' W_2' W_1 T,$$

(3.10) $$\partial W' V^{-1} W / \partial W = V^{-1} W + V'^{-1} W,$$

(3.11) $$\partial W' V^{-1} W / \partial V = -V'^{-1} W W' V'^{-1} .$$

Both equations (3.10) and (3.11) use the chain rule, while (3.11) then uses (3.5).

The following propositions are necessary, and are basic applications of the chain rule, the product rule, and equations (3.5) - (3.11).

Proposition 1. For $S(p \times 1)$

(3.12) $$\partial (X-S) \rho_1 \big(d(X-S,V)\big) / \partial S' \Big|_{S=0}$$
$$= -\rho_1 \big(d(X,V)\big) \otimes I_p - \rho_3 \big(d(X,V)\big) XX' (V'^{-1} + V^{-1}).$$

Proof: By the product rule and then the chain rule, (3.12) becomes

$$-\rho_1(d(X_0,V))\ I_p + X\left[\rho_3(d(X,V))*(\partial X'V^{-1}X/\partial X*\partial(X-S)/\partial S'\Big|_{S=0})\right].$$

Since $\partial(X-S)/\partial S' = I_p$, one applies (3.6) and (3.10). □

By the chain rule and (3.11) we obtain

Proposition 2. For $V(p\times p)$,

$$\partial\rho_1(d(X,V))/\partial V = -\rho_3(d(X,V))(V'^{-1}XX'V'^{-1}).$$

Proposition 3. For $S(p\times 1)$,

$$(3.13)\quad \partial(X-S)\rho_2(d^2(X-S,V))(X-S)'/\partial S'\Big|_{S=0}(I_p\otimes T)$$
$$= -\rho_2(d^2(X,V))(TX' + XT')$$
$$- \rho_2'(d^2(X,V))XX'(V'^{-1} + V^{-1})TX'.$$

Proof: By the chain rule and product rule as in Proposition 1, (3.13) becomes

$$-\rho_2(d^2(X,V))(X'\otimes I_p+XE_{(1,p)})(I_p\otimes T)$$
$$-\rho_2'(d^2(X,V))XX'(V'^{-1}+V^{-1})(X'\otimes I_p)(I_p\otimes T).$$

Application of (3.7) and (3.8) yields the result. □

Another application of the chain rule and (3.11) gives

Proposition 4. For $V(p\times p)$,

$$\partial\rho_2(d^2(X,V))/\partial V = -\rho_2'(d^2(X,V))V'^{-1}XX'V'^{-1}.$$

At this point it is useful to recall that Maronna has shown that V_0 is symmetric, thus simplifying the application of the preceeding propositions.

Proof of Theorem 2: Starting with the first defining equation, By Taylor expansions, (3.1), and the Law of the Iterated Logarithm,

$$\begin{aligned}
0 &= n^{-1}\textstyle\sum_1^n (X_i - T_n)\rho_1(d(X_i - T_n), V_n) \\
&= n^{-1}\textstyle\sum_1^n X_i\rho_1(d(X_i, V_n)) \\
&\quad + n^{-1}\textstyle\sum_1^n \{\partial(X_i - S)\rho_1(d(X_i - S, V_n))/\partial S'\big|_{S=0}\}T_n + O(n^{-1}\log_2 n) \\
&= n^{-1}\textstyle\sum_1^n X_i\rho_1(d(X_i, V_0)) \\
&\quad + n^{-1}\textstyle\sum_1^n X_i\left(\partial\rho_1(d(X_i, V_0))/\partial V_0\right)^{*\prime}(V_n - V_0)^* \\
&\quad + n^{-1}\textstyle\sum_1^n \{\partial(X_i - S)\rho_1(d(X_i - S, V_0))/\partial S'\big|_{S=0}\}T_n + O(n^{-1}\log_2 n) \\
&= n^{-1}\textstyle\sum_1^n X_i\rho_1(d(X_i, V_o)) \\
&\quad + E_1\{X_1\left(\partial\rho_1(d(X_1, V_0))/\partial V_0\right)^{*\prime}\}(V_n - V_0)^* \\
&\quad + E_1\{\partial(X_1 - S)\rho_1(d(X_1 - S, V_0))/\partial S'\big|_{S=0}\}T_n + O(n^{-1}\log_2 n),
\end{aligned}$$

the last following from the Law of the Iterated Logarithm. Propositions 1 and 2 then give (3.2). To obtain (3.3), note that

$$
\begin{aligned}
V_n^* &= n^{-1}\textstyle\sum_1^n\left[(X_i-T_n)\rho_2(d^2(X_i-T_n,V_n))(X_i-T_n)'\right]^* \\
&= n^{-1}\textstyle\sum_1^n(X_iX_i')^*\rho_2(d^2(X_i,V_n)) \\
&\quad + n^{-1}\textstyle\sum_1^n\{\partial(X_i-S)\rho_2(d^2(X_i-S,V_n))(X_i-S)'/\partial S'\Big|_{S=0}\ (I_p\otimes T_n)\}^* \\
&\quad + O(n^{-1}\log_2 n) \\
&= n^{-1}\textstyle\sum_1^n(X_iX_i')^*\rho_2(d^2(X_i,V_0)) \\
&\quad + n^{-1}\textstyle\sum_1^n(X_iX_i')^*(\partial\rho_2(X_i,V_0)/\partial V_0)^{*\prime}(V_n-V_0)^* \\
&\quad + n^{-1}\textstyle\sum_1^n\{\partial(X_i-S)\rho_2(d^2(X_i-S,V_0))(X_i-S)'/\partial S'\Big|_{S=0}\ (I_p\otimes T_n)\}^* \\
&\quad + O(n^{-1}\log_2 n).
\end{aligned}
$$

By the Law of the Iterated Logarithm, this gives

$$
\begin{aligned}
(V_n-V_0)^* &= n^{-1}\textstyle\sum_1^n\{X_i\rho_2(d^2(X_i,V_0))X_i' - V_0\}^* \\
&\quad + E_1\left[(X_1X_1')^*\left(\partial\rho_2(d^2{*}X_1,V_0))/\partial V_0\right)\right]^{*\prime}(V_n - V_0)^* \\
&\quad + \left\{E_1\{\partial(X_1-S)\rho_2(d^2(X_1-S,V_0))(X_1-S)'/\partial S'\Big|_{S=0}\}(I_p\ T_n)\right\}^* \\
&\quad + O(n^{-1}\log_2 n).
\end{aligned}
$$

Thus, (3.3) follows now from Propositions 3 and 4. Finally, to obtain (3.4), one plugs in the result (3.3) into (3.2), with manipulations following from the identify (3.9). For example

$$
\begin{aligned}
&E_1\{\rho_3(d(X_1,V_0))X_1(V_0^{-1}X_1'V_0^{-1})^{*\prime}\}H_2^{-1}E_0\{\rho_2(d^2(X_0,V_0))(T_nX_0' + X_0T_n')^*\} \\
&\qquad = E_{01}\{\rho_3(d(X_1V_0))\rho_2(d^2(X_0,V_0))X_1X_0'(W(X_1)+W(X_1)')\}T_n \quad . \quad \square
\end{aligned}
$$

Note that all the results would have been $O_p(n^{-1})$ if it were merely known that $n^{\frac{1}{2}}(T_n - T(P))$ had limit distributions. This we assume in the following Corollaries, although strong consistency of (T_n,V_n) would lead to almost sure results.

Corollary 1. If H_1 and H_2 are invertible, P is symmetric, and both $n^{1/2}(t_n - T(F))$ and $n^{1/2}(V_n - V_0(F))$ have limit distributions, then setting $T(F) = 0$ and $V_0(F) = V_0$,

$$T_n = H_1^{-1} n^{-1} \sum_1^n X_i \rho_1(d(X_i, V_0)) + O_p(n^{-1}) \tag{3.14}$$

$$V_n - V_0)^* = H_2^{-1} n^{-1} \sum_1^n (X_i \rho_2(d^2(X_i, V_0)) X_i' - V_0)^* + O_p(n^{-1}) \ . \tag{3.15}$$

If P is of radial type, one shows that H_1^{-1} exists and that the limit distribution of T_n is as given by Maronna. It might be of interest to note that (3.14) and (3.15) are considerably neater than (3.2) - (3.4). The following is a representation for univariate Huber Proposal 2.

Corollary 2. For p = 1, define $\psi_i(x) = x\rho_i(|x|)$ and for simplicity, set $T(F) = 0$, $V_0(F) = 1$. Then

$$(E\psi_1'(X_1) - E\{X_1\psi_1'(X_1)\}E\{X_1\psi_2'(X_1^2)\}/E\{X_1^2\psi_2'(X_1^2)\})T_n$$
$$= n^{-1}\sum_1^n \psi_1(X_i) - \tfrac{1}{2}(E\{X_1\psi_1'(X_1)\}/E\{X_1^2\psi_2'(X_1^2)\})n^{-1/2}\sum_1^n\{\psi_2(X_1^2)-1\} + O_p(n^{-1}) \ .$$

REFERENCES

1. Bickel, P.J. (1975). One step Huber estimates in the linear model. *J. Amer. Statist. Assoc. 70,* 428-434.

2. Carroll, R.J. (1977). On almost sure expansions for M-estimates. (To appear in *Ann. Statist.*)

3. Collins, J.R. (1976). Robust estimation of a location parameter in the presence of asymmetry. *Ann. Statist. 4,* 68-85.

4. Govindarajulu, Z. (1975). *Sequential Statistical Procedures.* Academic Press, New York.

5. Huber, P.J. (1967). The behavior of maximum likelihood estimates under non-standard conditions. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.*, 1, 221-233. Univ. of California Press.

6. MacRae, E.C. (1974). Matrix derivatives with an application to an adaptive linear decision problem. *Ann. Statist. 2,* 337-346.

7. Maronna, R. A. (1976). Robust M-estimates of multivariate location and scatter. *Ann. Statist. 4,* 51-67.

8. De Waal, D. J. (1975). *Parametric Multivariate Analysis, Part* 1. A series of lecture notes available from the Department of Statistics, University of North Carolina at Chapel Hill.