数理解析研究所講究録
第645巻1988年1-19

拡散過程におけるロバスト推定について

阪大基礎工　吉田朋広

1．はじめに．

この論文の目的は定常拡散過程のロバスト推定について考察することである．そのために，ここではM-推定の方法を用いる．つまり，推定量はある推定方程式の解としてあたえられる.M-推定量の漸近的な挙動はセクション3で述べられる．

拡散過程の定常分布が汚染されるとそれはM-推定量のバイアスに悪影響を及ぼす．セクション4で無限小の汚染の推定量に対する影響の尺度であるinfluence function (IF)を計算する．このとき，M-推定量の漸近分散はIFの2乗積分ではなくIFのある種の積分作用素による像の2乗積分になっている．IFの一様有界性のもとで最適なM-推定の存在を示すが，この理由で Künsch [10] のようにHuber型の最適解の陽な表現は困難なようである．そこで，セクション5では近似的に最適なM-推定を与える実際的方法を示す．

いっぽう，ロバスト性の評価に異なった規準をとれば，別の最適化問題も考えられる．セクション6ではIFの大きさの尺度として積分形のノルムをとる．このとき，最適ロバストM-推定は2階微分方程式を解くことによって与えられることを示す．最後に，具体例をあげる．

2．モデルと最尤推定量．

次の確率微分方程式をかんがえる：

(2.1)　$dX(t)=\sigma(X(t))dW(t)+f(X(t),\theta)dt$

　　　$X(0)=x(0)$.

ここで，θはパラメータ，$\theta\in\Theta\subset R$，$f(x,\theta)$と$\sigma(x)$は連続微分可能な関数，$\sigma(x)$は正，$W(t)$は標準Wiener過程とする．確率微分方程式の解の存在のための十分条件は係数のLipschitz連続性と線形増大性である．(Ikeda and Watanabe [8] を見よ．) 我々は(2.1)を満足する拡散過程を考える．

さらに，拡散過程のうちエルゴード性を持つものを考える．相空間の境界$\pm\infty$はinaccessibleとする．

(2.2)　$B(x,\theta)=\int_0^x 2\sigma^{-2}(y)f(y,\theta)dy$,

(2.3)　$m(x,\theta)=\int_0^x 2\sigma^{-2}(y)\exp B(y,\theta)dy$,

とおくと，$m(\infty,\theta)-m(-\infty,\theta)<\infty$のとき，(2.1)で定義される拡散過程はエルゴード性をもち，その定常分布は$\nu(x,\theta)=[m(x,\theta)-m(-\infty,\theta)]/[m(\infty,\theta)-m(-\infty,\theta)]$で与えられる．$\mu(x,\theta)=\nu(dx,\theta)/dx$とおく．　詳しくは，Gihman and Skorohod [5]を見よ．

観測$(X(t);0\leqq t\leqq T)$にもとずく対数尤度関数は

(2.4)

$$\Lambda(T,\theta)=\int_0^T f(X(t),\theta)\underline{\sigma}^{-2}(X(t))dX(t)-\frac{1}{2}\int_0^T f^2(X(t),\theta)\sigma^{-2}(X(t))dt$$

で与えられる．(Liptser and Shiryayev [16]参照．) $H(x,\theta):=B(x,\theta)/2$ と置けば，伊藤の公式によって，

(2.5)　$\Lambda(T,\theta)=H(X(T),\theta)-H(X(0),\theta)-\int_0^T h(X(t),\theta)dt$,

ここで，

(2.6)　$h(x,\theta)=[f(x,\theta)\sigma^{-1}(x)]^2/2+\sigma^2(x)\partial/\partial x[f(x,\theta)\sigma^{-2}(x)]/2$

となる．ゆえに，正則条件のもとで，最尤推定量(MLE)は方程式

(2.7)　$\dot{\Lambda}(T,\theta)=\dot{H}(X(T),\theta)-\dot{H}(X(0),\theta)-\int_0^T L(\theta)\dot{H}(X(t),\theta)dt=0$,

の解である．ここで，ドット"・"はパラメータに関する微分を表し，$L(\theta)$は(2.1)で定義される拡散過程の生成作用素である：

(2.8)　$L(\theta)=f(x,\theta)D+\frac{1}{2}\sigma^2(x)D^2$, $D=\frac{\partial}{\partial x}$

(2.9)　$L(\theta)\dot{H}(x,\theta)=\dot{h}=f\dot{f}\sigma^{-2}+\frac{1}{2}\sigma^2 D[\dot{f}\sigma^{-2}]$.

真のパラメータがθのとき$\dot{\Lambda}(T,\theta)$はマルチンゲールである．拡散過程におけるMLEの一致性，漸近正規性，有効性については，Kutoyants [11], [12], [13], Lånska [14], Rao and Rubin [20], McKeague [18] を見よ．

３．M-推定量とその漸近性質．

実際の場面においては，厳密な意味で，設定したモデル(2.1)から発生したデータを得ることはほとんど期待できないであろう．つまり，ある種のノイズによってデータが汚されていたり，真のモデルがパラメトリックモデルに含まれていなかったりすることがあるだろう．ガウス型モデルが合理的であって，また，扱いやすくても，実際のデータが裾の重い分布をもったノイズで汚染されていれば，そのモデルで推論することは良いとは限らない．そこで，我々はロバストな推定量を構成したい．

ロバスト推定の問題は，おもに，独立同一分布にたいして研究されてきた．これについては Hampel et al. [7] に詳しく述べられている． 従属性のあるモデル,とくに，時系列モデルのロバストなパラメータ推定に関しては，最近盛んに研究されてきている．Denby and Martin [3]，Bustos [1]，Künsch [10]，Martin and Yohai [17]，Bustos and Yohai [2] 等を見よ．これらの論文でM-推定，一般化M-推定と呼ばれている推定法に対応して，

定義 3.1．関数$a(x,\theta)$と$A(x,\theta)$に対して，推定方程式

$$(3.1)\qquad M(T,\theta):=A(X(T),\theta)-A(X(0),\theta)-\int_0^T a(X(t),\theta)dt=0$$

で定義される推定量$\hat{\theta}(T)$をa，Aに対応するM-推定量とよぶ．□

最尤推定量は

$$(3.2)\qquad a=L(\theta)\dot{H}=\dot{h},\ A=\dot{H}$$

に対応するM-推定量である．我々の推定方程式(3.1)は Lánska [14] のコントラスト関数の微分と同じであって，M-推定量の漸近性質も同様であるが，ロバストな推定量を構成するにはより広い関数のクラス，つまり，Sobolev型空間，を扱う方が便利である．特に，ここでは，$x\to a(x,\theta)$の連続性を仮定しない．そこで，我々のM-推定量の漸近的性質を簡単に述べよう．

この論文においては$\int\sigma^2(x)\mu(x,\theta)dx<\infty\ (\theta\in\Theta)$を仮定する．　さらに，$F=\{A;DA\in L^2(\sigma^2\mu(\cdot,\theta)),\ L(\theta)A\in L^2(\mu(\cdot,\theta)),\ \theta\in\Theta\}$とする．ここで，xに関する微分はSchwartzの超関数の意味で，$L^2(\sigma^2\mu(\cdot,\theta))$，$L^2(\mu(\cdot,\theta))$

は測度 $\sigma^2\mu(x,\theta)dx$，$\mu(x,\theta)dx$に関するL^2空間を表す．また，R上の積分は積分区間を省略する．簡単のため，定常拡散過程を考え，以下の仮定をする．

(I) $A(\cdot,\theta)\in F$，$\theta\in\Theta$．

(II-1) $\theta\in\Theta$，$\theta'\in\mathrm{Int}\Theta$に対して，$a(\cdot,\theta)\in L^1(\mu(\cdot,\theta'))$．

(II-2) $\theta'\in\mathrm{Int}\Theta$に対して，

$$\frac{\partial}{\partial\theta}\int a(x,\theta)\mu(x,\theta')dx\Big|_{\theta=\theta'}=\int \dot{a}(x,\theta')\mu(x,\theta')dx\neq 0.$$

(III) $\theta'\in\mathrm{Int}\Theta$に対して，$P_{\theta'}$のもとで確率1で，$\theta\to M(T,\theta)$，は$\theta'$で連続，$T>0$．

(IV) $\theta\in\mathrm{Int}\Theta$にたいして，$\dot{a}(\cdot,\theta)\in L^1(\mu(\cdot,\theta))$，さらに，

$$\dot{M}(T,\theta)/T\to-\int\dot{a}(x,\theta)\mu(x,\theta)dx \text{ in } P_\theta$$

が連続的に成り立つ．つまり，$\dot{M}$ のなかのθをθに収束する任意の列s(T)に置き換えてもうえの収束が成り立つ．

(V) $DG(\cdot,\theta)\in L^2(\sigma^2\mu(\cdot,\theta))$，$\theta\in\Theta$，ここで

$$G(x,\theta)=-\int_0^x\exp(-B(y,\theta))dy\int_y^\infty 2a(u,\theta)\sigma^{-2}(u)\exp(B(u,\theta))du. \tag{3.3}$$

もしA=G ならば，$L(\theta)G(\cdot,\theta)=a(\cdot,\theta)$であるので，$P_\theta$ のもとで，$M(T,\theta)$は局所マルチンゲールになる．

注意 3.1．後に続く2つの定理のためにはFを次のようにとれば，十分である：

$$F=\{A;DA\in L^2(\sigma^2\mu(\cdot,\theta)),\ L(\theta)A\in L^1(\mu(\cdot,\theta)),\ \theta\in\Theta\}.\square$$

定理 3.1．条件(I)，(II-1,2)，(III)，(V) を仮定する．$\theta\in\mathrm{Int}\Theta$が真ならば，一致性を持つM-推定量$\hat{\theta}(T)$が存在する．□

証明の前に補題を用意する．

補題 3.1．$\theta'\in\mathrm{Int}\Theta$が真のとき，$\theta\in\Theta$に対して

$$[A(X(T),\theta)-A(X(0),\theta)]/T\to 0.\quad \text{a.s.}(P_{\theta'})\quad T\to\infty.$$

証明．拡張された伊藤の公式 (Krylov [9])によって，

(3.4) $[A(X(T),\theta)-A(X(0),\theta)]/T$

$$=\frac{1}{T}\int_0^T L(\theta')A(X(t),\theta)dt+\frac{1}{T}\int_0^T DA(X(t),\theta)\sigma(X(t))dW(t).$$

補題の証明のためには，(3.4)の右辺の第一項が零に確率1で零に収束することを示せばよい．実際，第二項はLepingle [15] のマルチンゲールに対する強大数の法則で零に確率1で収束するから．　さらに，エルゴード性によって，$\int L(\theta')A(x,\theta)\mu(x,\theta')dx=0$を示せば十分である．DAに対して

$$DA(x,\theta)=\exp(-B(x,\theta'))[\int_{-\infty}^{x}2\sigma^{-2}L'(\theta')DA(y,\theta)\exp B(y,\theta')dy+c],$$

ここで，cは定数，$L'(\theta')=f(\cdot,\theta')+\frac{1}{2}\sigma^2(\cdot)D$.

$$k=\int 2\sigma^{-2}L'(\theta')DA(y,\theta)\cdot\exp B(y,\theta')dy+c$$

と置く．　$k\neq 0$とすれば，$x\to\infty$のとき$\exp B(x,\theta')\cdot DA(x,\theta)\to k$だから，$\sigma^2(x)\mu(x,\theta')DA(x,\theta)$のLebesgue測度に関する可積分性に反する．（この可積分性は仮定とSchwarzの不等式からわかる．）ゆえに，k=0.同様に，c=0も示せて，結局，$\int L'(\theta')DA(x,\theta)\mu(x,\theta')dx=0$ が言えて証明が終わる．□

定理 3.1 の証明．　十分小さな任意の$\delta>0$に対して，　補題3.1によって，$M(T,\theta\pm\delta)/T$は$-\int a(x,\theta\pm\delta)\mu(x,\theta)dx$に確率1($P_\theta$)で収束する．

$\int a(x,\theta)\mu(x,\theta)dx=0$となることは補題3.1の証明と同じ議論でわかるから，仮定(II-2)によって，十分大きなTに対して$M(T,\theta\pm\delta)$の符号は異なる．従って，(III)より$\hat\theta(T)$が$(\theta-\delta,\theta+\delta)$に存在することがわかる．□

定理 3.2.　条件(I)-(V)のもとで，$\theta\in\mathrm{Int}\Theta$が真のとき，

$$\mathcal{L}\{\sqrt{T}(\hat\theta(T)-\theta)\}\to N(0,\Sigma),\quad T\to\infty.$$

ここで，

$\Sigma=\Delta/U^2$, $\Delta=\int[DG(x,\theta)\sigma(x)]^2\mu(x,\theta)dx$, $U=U(\theta)=-\int\dot a(x,\theta)\mu(x,\theta)dx$.

証明．　テイラー展開によって，

$$-T^{-1/2}M(T,\theta)=T^{-1}\dot M(T,\tilde\theta)T^{1/2}(\hat\theta(T)-\theta).$$

ここで，$\tilde\theta$は$\hat\theta(T)$とθの間にある．条件(IV)によって，

(3.5) $UT^{1/2}(\hat\theta(T)-\theta)+T^{-1/2}M(T,\theta)\to 0$ in P_θ.

いっぽう,

$$T^{-1/2}M(T,\theta)=T^{-1/2}[A(X(T),\theta)-A(X(0),\theta)]$$
$$-T^{-1/2}[G(X(T),\theta)-G(X(0),\theta)]$$
$$(3.6)\qquad +T^{-1/2}\int_0^T DG(X(t),\theta)\cdot\sigma(X(t))dW(t).$$

式(3.6)の右辺のはじめの2つの項は定常性によって零に確率収束することがわかる．第3項はマルチンゲールにたいする中心極限定理(例えば，Feigin [4])をつかえば$N(0,\Delta)$に分布収束することがわかる．(3.5)と(3.6)より定理の結果がしたがう．□

4．Influence function とロバスト性．

汚染の影響を受けているかも知れない観測$X(t)$の定常分布をκとしよう．幾らかの仮定のもとで,

(4.1) $\quad M(T,\theta)/T\to -\int a(x,\theta)d\kappa$ a.s.

となることがわかる．(4.1)の右辺が零になるようなθが一意であるとして，それを$Y(\kappa)$とかく．　さらに，(3.1)の収束が$C(\Theta)$の一様位相の意味でなりたてば$\hat\theta(T)$は$Y(\kappa)$に収束する．このようなaとκに対して,

(4.2) $\quad \int a(x,Y(\kappa))d\kappa=0.$

ここで，κは$\{\nu(\cdot,\theta)\}$に属さなくともよいことに注意せよ．Fisherの一致性は$Y(\nu(\cdot,\theta))=\theta\ (\theta\in\Theta)$と書ける．

もし，$X(t)$の定常分布$\nu(\cdot,\theta)$が$(1-t)\nu(\cdot,\theta)+t\kappa$（$\kappa$はR上の確率分布)に変化したとすると，(4.2)より,

(4.3) $\quad \int a(x,Y((1-t)\nu(\cdot,\theta)+t\kappa))d((1-t)\nu(\cdot,\theta)+t\kappa)=0$

となる．tに関してt=0で微分すると,

(4.4) $$Y'(\kappa,a,\nu(\cdot,\theta)):=(\partial/\partial t)Y((1-t)\nu(\cdot,\theta)+t\kappa)|_{t=0}$$
$$=U^{-1}\int a(x,\theta)d\kappa,$$

ここで,

(4.5) $U=U(\theta)=-\int \dot{a}(x,\theta)\mu(x,\theta)dx.$

極限的な$\nu=\delta_x$の場合，

(4.6) $Y'(x,a,\nu(\cdot,\theta)):=Y'(\delta_X,a,\nu(\cdot,\theta))=U^{-1}a(x,\theta)$

とおき，これをM-推定の$\nu(\cdot,\theta)$におけるinfluence function(IF)とよぶ．IFはXの定常分布の変化の推定量への影響の度合を測るひとつの尺度である．

Fisherの一致性より，

$$\int a(x,\theta)\mu(x,\theta)dx=0$$

であって，パラメータに関して微分することによって，

(4.7) $U^{-1}\int a(x,\theta)\dot{\mu}(x,\theta)dx=1$

をえる．

MLEに対するIFはしばしば非有界であるので，(4.4)からわかるように，定常分布の汚染は，特にそれが大きな$|x|$に対しては，推定量に与える影響が大きい．そこで，適当な関数a，Aを選んでM-推定の手法でロバストな推定量を構成したい．このとき，IFは何等かの意味で小さくなるようにし，同時に，推定量の漸近分散が最小あるいはMLEと比べてあまり大きくならないように構成するのが好ましい．IFの評価規準として，まず，Hampel [6] にしたがって，sup-ノルムつまりgross error sensitivityを採用する．

セクション3で見たように，M-推定量の漸近分散は本質的にAには依らずaにのみ依存する．したがって，以後，Aのかわりに，aに依存するGをとる．そこで第一段階として，我々の最適化問題は次の様になる：

(P1)以下の条件のもとで，

$$\int [\sigma(x)DG/U]^2\mu(x,\theta)dx$$

を最小にせよ：

(4.8) $\int L(\theta)G/U\cdot\mu(x,\theta)dx=0,$

(4.9) $\mathrm{ess.sup}|L(\theta)G/U|\leq c(\theta),$

(4.10) $\int L(\theta)G/U\cdot\dot{\mu}(x,\theta)dx=1,$

ここで，$c(\theta)$は定数で，

(4.11) $G\in F(\theta):=\{G;DG\in L^2(\sigma^2\mu(\cdot,\theta)),\ L(\theta)G\in L^2(\mu(\cdot,\theta))\}$.

$L(\theta)G=a$であることに注意せよ．漸近分散はIFの2乗積分ではない．

最適化問題(P1)は$\xi=DG/U$と置くことによって，次の問題(P2)と同値である．

(P2)以下の条件のもとで，

$$\int \xi^2\sigma^2\mu(x,\theta)dx$$

を最小にせよ：

(4.8)′ $\int L'(\theta)\xi\cdot\mu(x,\theta)dx=0$,

(4.9)′ $\text{ess.sup}|L'(\theta)\xi|\leqq c(\theta)$,

(4.10)′ $\int L'(\theta)\xi\cdot\dot{\mu}(x,\theta)dx=1$,

ここで，$c(\theta)$は定数で，$L'(\theta)=f(\cdot,\theta)+\frac{1}{2}\sigma^2(\cdot)D$,

(4.11)′ $\xi\in F'(\theta):=\{\xi:\xi\in L^2(\sigma^2\mu(\cdot,\theta)),\ L'(\theta)\xi\in L^2(\mu(\cdot,\theta))\}$.

$\dot{\mu}(\cdot,\theta)/\mu(\cdot,\theta)\in L^2(\mu(\cdot,\theta))$とし，CをF′($\theta$)の元で条件(4.8)′-(4.10)′を満たすものの全体とする．

定理 4.1．Cが空でなければ(P2)の最適解がCのなかに存在する．□

定理の証明のまえにいくつかの準備をする．ノルム

$$|\xi|_{F'}^2=\int \xi^2\sigma^2\mu(x,\theta)dx+\int (L'(\theta)\xi)^2\mu(x,\theta)dx$$

を$F':=F'(\theta)$に与えれば，F′はHilbert空間になる．さらに，

$$|\xi|_2^2=\int \xi^2\sigma^2\mu(x,\theta)dx,$$

$$L^2=L^2(\sigma^2\mu(\cdot,\theta)).$$

とかく．

補題 4.1．($\xi(n);n\geqq 1$)をCのなかの任意の列とする．もし，$\xi(n)$が$\xi\in F'$にF′で弱収束するならば，(1)$\text{ess.sup}|L'(\theta)\xi|\leqq c(\theta)$，(2)($\xi(n);n\geqq 1$)の部分列($\xi(n')$)が存在して$\xi(n')$は$\xi$に$L^2$で弱収束し，$L'(\theta)\xi(n')$は$L'(\theta)\xi$に$L^2(\mu(\cdot,\theta))$で弱収束する．

証明. θを固定し省略する. $(\xi(n))$は弱有界だから, Banach-Steinhausの定理によって, $(\xi(n))$はF'において有界である. 定義から$|\cdot|_2 \leqq |\cdot|_{F'}$だから$(\xi(n))$は$L^2$においても有界である. したがって,$(\xi(n);n\geqq 1)$の部分列$(\xi(n'))$が存在して$\xi(n')$はある$\xi' \in L^2$に弱収束する. 同様にして, $(L'\xi(n))$の$L^2(\mu)$における有界性から$L'\xi(n')$はある$\eta \in L^2(\mu)$に$L^2(\mu)$で弱収束すると仮定してよい. $\mathcal{D}$をR上のコンパクトな台をもつ滑らかな関数全体の空間とする. Schwartzの意味の超関数の空間を$\mathcal{D}'$とかき, $\mathcal{D}' \times \mathcal{D}$上の双線形形式を$<\ ,\ >$で表す. $(\xi(n'))$と$(L'\xi(n'))$の弱収束から, 任意の$\psi \in \mathcal{D}$に対して, $<L'\xi', \psi> = <\eta, \psi>$がわかる. したがって, $L'\xi' = \eta$, ゆえに, $(\xi(n'))$はξ'にF'で弱収束する. 弱収束極限の一意性より$\xi' = \xi$が言える.

任意の$\psi \in \mathcal{D}$に対して, $L'\xi(n')$の$L'\xi$への超関数の意味での収束から, $|<|L'\xi|, \psi>| \leqq c|\psi|_{L^1(R)}$が言えて, $\mathrm{ess.sup}|L'\xi| \leqq c$もわかる. □

補題 4.2. CはF'で弱閉である.

証明. $\xi(n) \in C$がξにF'で弱収束するとせよ. 補題4.1より, $|L'(\theta)\xi| \leqq c$.また, $(\xi(n);n\geqq 1)$の部分列$(\xi(n'))$が存在して, $L'(\theta)\xi(n')$は$L'(\theta)\xi$に$L^2(\mu(\cdot,\theta))$で弱収束する. 1と$\dot{\mu}(\cdot,\theta)/\mu(\cdot,\theta)$は$L^2(\mu(\cdot,\theta))$の元だから, (4.8)', (4.10)'は容易にわかる. したがって, $\xi \in C$. □

定理 4.1 の証明. $k = \inf\{|\xi|_2 ; \xi \in C\}$とする. このとき, Cの元の列$(\xi(n))$で$n \to \infty$のとき$|\xi(n)|_2 \to k$となるものが存在する. Cでは$|L'\xi| \leqq c$だから, $|\xi(n)|_{F'}$は有界で, $(\xi(n))$の部分列$(\xi(n'))$が存在して$\xi(n')$はξにF'で弱収束する. このとき, 補題4.2から, $\xi \in C$がいえる. また, 補題4.1によれば, $(\xi(n'))$の部分列$(\xi(n''))$が存在して, $(\xi(n''))$はξにL^2で弱収束する. Hilbert空間の理論から, $|\xi|_2 \leqq \liminf|\xi(n'')|_2 = k$となることがわかる. ゆえに, $|\xi|_2 = k$となって, このξが最適解である. □

注意. 4.1. 補題3.1の証明からわかるように条件$\xi \in L^2(\sigma^2\mu(\cdot,\theta))$は一致性の条件(4,8)'を導くので(4,8)'は必要ない. しかしながら, 一致性を明確にするためにこの条件も付け加えた. □

5．ロバストな推定を構成する簡単な方法.

最適化問題(P2)は線形制約のもとでの最小化問題でであって，その数値解析的なアプローチは多くの著者によって研究されている.（Powell [19].） しかし，比較的よい結果を与える簡単な別の方法がある．これはHampelの最適解の作り方の類似である.

$v(+),v(-)$を$[0,\infty]$から選んだ定数とする．$v(+),v(-)$は一般にθに依存してもよい．以後，記号の簡単のためしばしばθを省略する．つぎのように関数ϕを決める：

$$(5.1)\qquad \phi=\begin{cases} l_-(x), & x<-v(-) \\ \lambda(=-\dot{f}\sigma^{-2}), & -v(-)\leqq x\leqq v(+) \\ l_+(x), & v(+)<x \end{cases}$$

ここで，

$$(5.2)\qquad l_-(x)=\exp(-B(x,\theta))\int_{-\infty}^{x} 2k_-\sigma^{-2}(y)\exp(B(y,\theta))dy,$$

$$(5.3)\qquad l_+(x)=-\exp(-B(x,\theta))\int_{x}^{\infty} 2k_+\sigma^{-2}(y)\exp(B(y,\theta))dy,$$

さらに，定数k_-とk_+は次の式で定義される：

$$(5.4)\qquad \lambda(-v(-))=l_-(-v(-))$$

$$(5.5)\qquad \lambda(v(+))=l_+(v(+)).$$

このとき，ϕは連続で，$L'(\theta)\phi$は$x<-v(-)$，$-v(-)\leqq x\leqq v(+)$，$v(+)<x$上でそれぞれ，k_-，$-\dot{h}$，k_+に等しい.

$l_+ 1_{[0,\infty)}$，$l_- 1_{(-\infty,0]}$，λ，$1\in L^2(\sigma^2\mu(\cdot,\theta))$とする．このとき，

$$(5.6)\qquad S(\phi)=\int\phi\lambda\sigma^2\mu\,dx$$

$$=\int_{-v(-)}^{v(+)}(\dot{f}/\sigma)^2\mu dx-\int_{v(+)}^{\infty}l_+\dot{f}\mu dx-\int_{-\infty}^{-v(-)}l_-\dot{f}\mu dx,$$

(5.7) $$Q(\phi)=\int\phi^2\sigma^2\mu dx$$

$$=\int_{-v(-)}^{v(+)}(\dot{f}/\sigma)^2\mu dx+\int_{v(+)}^{\infty}l_+^2\sigma^2\mu dx+\int_{-\infty}^{-v(-)}l_-^2\sigma^2\mu dx.$$

我々の近似的に最適なM-推定量の候補は

(5.8) $$\xi=DG/U=\phi/S(\phi)$$

に対応する．このIFは

(5.9) $$L'(\theta)\xi=\begin{cases}k_-/S(\phi), & x<-v(-)\\ -\dot{h}/S(\phi), & -v(-)\leqq x\leqq v(+)\\ k_+/S(\phi), & v(+)<x\end{cases}$$

となる．(5.8)で定義されたξに対して，

命題 5.1. もし，$x\to\pm\infty$のとき$\phi\sigma^2\dot{\mu}=0$ならば，(4.8)′-(4.11)′が

(5.10) $$c(\theta)=\sup\{|k_+|,|k_-|,|\dot{h}(x,\theta)|;-v(-)\leqq x\leqq v(+)\}/|S(\theta)|$$

に対して成立する．さらに$\int\xi^2\sigma^2\mu dx$(ξに対応するM-推定量の漸近分散)は$Q(\phi)/S(\phi)^2$で，これはFisher情報量の逆数$[\int(\dot{f}/\sigma)^2\mu dx]^{-1}$に$v(+),v(-)\to\infty$のとき収束する．□

注意 5.1. この極限はMLEの漸近分散に等しい．□

補題 5.1. $L'(\theta)^*\cdot\dot{\mu}\sigma^{-2}\mu^{-1}=-\dot{f}\sigma^{-2}=\lambda$，ここで，$L'(\theta)^*$ は$L'(\theta)$の共役作用素，つまり，

$$L'(\theta)^*=f(\cdot,\theta)-\tfrac{1}{2}D\sigma^2.$$

証明．(2.2)と(2.3)を使った簡単な計算による．□

命題 5.1の証明．(4.9)′と(4.11)′は容易．(4.8)′は注意4.1による．補題5.1と部分積分によって，

$$\int_{-N}^{N}L'(\theta)\phi\cdot\dot{\mu}dx=\int_{-N}^{N}\phi L'(\theta)^*\dot{\mu}dx+[\phi\sigma^2\dot{\mu}/2]_{-N}^{N}$$

$$=\int_{-N}^{N}\phi\lambda\sigma^2\mu dx+[\phi\sigma^2\dot{\mu}/2]_{-N}^{N}.$$

$N\to\infty$として，

$$\int L'(\theta)\phi\cdot\dot{\mu}dx=\int\phi\lambda\sigma^2\mu dx=S(\phi),$$

となって，(4.10)′をえる．最後の部分は(5.6)と(5.7)から明らかである．□

さらに一般に，定数k_-とk_+の代わりに，有界連続関数$k_-(x,\theta)$と$k_+(x,\theta)$で，(5.2),(5.3)で定義されるl_-,l_+に対して，(5.4),(5.5)が成り立つものをとれば，命題5.1と同じ結果が(5.10)をつぎのようにしてなりたつ：

$$c(\theta)=\sup[\{|k_-|;x<-v(-)\}\cup\{|k_+|;v(+)<x\}\cup\{|\dot{h}|;-v(-)\leqq x\leqq v(+)\}]/|S(\phi)|.$$

もし，$\mu(x,\theta)$が$|x|\to\infty$のとき指数関数的に減少し，$l_\pm$，$\dot{f}/\sigma$が多項式のオーダーならば，$v(+)$，$v(-)\to\infty$のとき$S(\phi)$，$Q(\phi)$は Fisherの情報量に急速に収束し，M-推定量はほとんど最小漸近分散を持つ．いっぽう，$f\dot{f}/\sigma^2$，$\dot{h}$が多項式のオーダーのとき，(5.4)，(5.5)から，$|k_-|$，$|k_+|$は多項式のオーダーだから，$c(\phi)$はあまり大きくならない．

最後に，例をあげる．　Xをつぎの確率微分方程式で定義される拡散過程(Ornstein-Uhlenbeck の拡散過程)とする：

$$dX(t)=-\theta X(t)dt+dW(t),\ X(0)=x(0),\ \theta>0.$$

このとき，

$$f(x,\theta)=-\theta x,\ \sigma(x)=1,$$

$$\mu(x,\theta)=(\theta/\pi)^{1/2}\exp(-\theta x^2),\ \lambda=\lambda(x,\theta)=-\dot{f}\sigma^{-2}=x,$$

$$h(x,\theta)=\theta^2x^2/2-\theta/2,\ \dot{h}=\theta x^2-1/2.$$

(5.1)-(5.8)で定義されたξに対して，

$$\xi=\begin{cases} l_-(x)/S(\phi)=\exp(\theta x^2)\int_{-\infty}^{x}2k_-\exp(-\theta y^2)dy\,/S(\phi), & x<-v(-)\\ x/S(\phi), & -v(-)\leqq x\leqq v(+)\\ l_+(x)/S(\phi)=-\exp(\theta x^2)\int_{x}^{\infty}2k_+\exp(-\theta y^2)dy/S(\phi), & v(+)<x,\end{cases}$$

ここで，k_-，k_+は

$$k_-=-v(-)[2\exp(\theta v(-)^2)\int_{v(-)}^{\infty}\exp(-\theta y^2)dy]^{-1},$$

$$k_+=-v(+)[2\exp(\theta v(+)^2)\int_{v(+)}^{\infty}\exp(-\theta y^2)dy]^{-1},$$

$$\begin{aligned}S(\phi)=&\int_{-v(-)}^{v(+)}x^2(\theta/\pi)^{1/2}\exp(-\theta x^2)dx\\&-\int_{v(+)}^{\infty}2k_+x(\theta/\pi)^{1/2}\int_{x}^{\infty}\exp(-\theta y^2)dydx\\&+\int_{-\infty}^{-v(-)}2k_-x(\theta/\pi)^{1/2}\int_{-x}^{\infty}\exp(-\theta y^2)dydx.\end{aligned}$$

ξのIFは

$$L'(\theta)\xi=\begin{cases}k_-/S(\phi), & x<-v(-)\\(-\theta x^2+1/2)/S(\phi), & -v(-)\leqq x\leqq v(+)\\k_+/S(\phi), & v(+)<x\end{cases}$$

となる．とくに，MLEに対して対応するものは

$$S(\phi_{MLE})=\int\lambda^2\mu dx=1/(2\theta),$$

$$\xi_{MLE}=\lambda/S(\phi_{MLE})=2\theta x,$$

となり，IFは

$$L'(\theta)\xi_{MLE}=-\dot{h}/S(\phi_{MLE})=-2\theta(\theta x^2-1/2)=-2\theta^2x^2+\theta.$$

このとき，MLEのIFは2次関数で定常分布の汚染は，とくに$|x|$が大きいところでのそれは，バイアスに大きな影響を与える．$\theta=1/2$のときMLEの漸近分散(最小分散)は1であるが，分点v(-)，v(+)と上で作ったM-推定量のIFの上限$c(\theta)$と漸近分散は次のページの表で与えられる．我々のM-推定量は比較的小さな漸近分散を持つことがわかる．

表1. 分点，IFの上限，漸近分散.

v(-)(=v(+))	c(θ)	漸近分散
0.50	1.88	1.68
1.00	1.30	1.26
1.50	1.75	1.09
2.00	2.45	1.02
2.50	3.43	1.00
3.00	4.68	1.00
3.50	6.18	1.00

注意 5.2. 1)我々のM-推定量のIFは一般に不連続である.
2)ξを作ったあとでこれがセクション3の条件を満足するかどうか検討しなければならない. 上の例ではθに依らないv(-), v(+)と$\Theta=[a,b]$($a>0,b>0$)に対して，これは容易にわかる. □

6. 他の規準による最適化問題.

セクション4と5ではgross error sensitivity， つまり，IFの大きさの評価としてsup-ノルムを考えた. 推定量のロバスト化のためにはIFを何等かの意味で小さくするべきだが，IFの評価として他の規準を取ることも可能である.

$w(x)$をRで定義されたC^1-級の正値関数とする. 以前と同じ記号を使えば，ξに対応するM-推定量のIFは$L'(\theta)\xi$に等しく，この大きさをノルム

$$(6.1)\qquad \int |L'(\theta)\xi|^2 w(x)dx$$

で評価することができる. 確率過程Xの定常分布の汚染のM-推定量への影響は

(4.4)で与えられるので，もし，$d\kappa(x)\leqq w(x)dx$ならば，

$$|Y'(\kappa, a, \nu(\cdot,\theta))|=|\int L'(\theta)\xi d\kappa(x)|\leqq\{\int|L'(\theta)\xi|^2 w(x)dx\}^{1/2}$$

となって，$w(x)dx$で抑えられる汚染にたいしてノルム(6.1)はその影響の上界を与えている．

ノルム(6.1)を小さくし，同時に漸近分散

(6.2) $\quad |\xi|_2^2=\int\xi^2\sigma^2\mu dx$

を小さくする様なξを選ぶのが好ましい．そこで，新しくノルム

(6.3) $\quad |\xi|_W^2=\int\xi^2\sigma^2\mu dx+\int|L'(\theta)\xi|^2 w(x)dx$

を定め，$|\cdot|_w$ -ノルムを最小にするξを探す．そのようなξを最適w-ロバストM-推定量とよぶ．MLEはある正則な推定量のクラスで最適0-ロバストM-推定量である．

すべての$\theta\in\Theta$にたいして$\int\lambda^2\sigma^2\mu dx<\infty$，$\int\sigma^2\mu dx<\infty$を仮定する．$F(W,\theta)$を$\mathcal{D}$のノルム$|\cdot|_w$に関する完備化とする．このとき，$(F(W,\theta),\ |\cdot|_w)$はHilbert空間になる．$C(w,\theta)=F(w,\theta)\cap\{\xi;\int\xi\lambda\sigma^2\mu dx=1\}$とおく．(ここに現れた積分は，補題5.1によって，条件(4.10)′に対応している．)$\xi\in F(W,\ \theta)$に対して一致性の条件が成り立つことに注意せよ．そこで，我々の考える最適化問題は

(P3)$|\xi|_w$を$C(w,\theta)$で最小にせよ．

後で出てくる不等式(6.5)から，線形汎関数$\xi\to\int\xi\lambda\sigma^2\mu dx$は$F(W,\theta)$上有界なることがわかる．したがって，$C(w,\theta)$はHilbert空間$F(W,\theta)$の中の凸閉集合であることがいえて，$|\xi_0|_w=\min\{|\xi|_w;C(w,\theta)\}$となる$\xi_0\in C(w,\theta)$の一意的な存在がわかる．

しかし，この最適解ξ_0は境界値問題を解くことによって直接関数形を求めることができる．つぎの2階微分方程式を考える：

(6.4) $\quad L'(\theta)^*[w(x)L'(\theta)\xi]+\sigma^2(x)\mu(x,\theta)\xi=\lambda(x,\theta)\sigma^2(x)\mu(x,\theta)$

ここで，$L'(\theta)$，$L'(\theta)^*$，$\mu(x,\theta)$等は以前と同じで，xに関する微分は超関数

の意味である.

定理 6.1. 微分方程式(6.4)の解が$F(W,\theta)$のなかに一意に存在する.

証明. つぎの評価が成り立つ:

$$(6.5)\qquad \left|\int \xi\lambda\sigma^2\mu dx\right| \leqq \left\{\int \lambda^2\sigma^2\mu dx\right\}^{1/2}\left\{\int \xi^2\sigma^2\mu dx\right\}^{1/2}$$
$$\leqq \left\{\int \lambda^2\sigma^2\mu dx\right\}^{1/2}|\xi|_W.$$

したがって, 線形汎関数$\xi\to\int\xi\lambda\sigma^2\mu dx$は$F(W,\theta)$上有界である. Rieszの定理によって, ある$\xi_1$が$F(W,\theta)$の中に一意に存在して, 任意の$\xi\in F(W,\theta)$に対して,

$$(6.6)\qquad \int \xi\xi_1\sigma^2\mu dx = \int (L'(\theta)\xi)(L'(\theta)\xi_1)wdx = \int \xi\lambda\sigma^2\mu dx.$$

とくに, $\xi\in\mathcal{D}$に対して,

$$\int \xi\{L'(\theta)^*[wL'(\theta)\xi_1]+\sigma^2\mu\xi_1\}dx = \int \xi\lambda\sigma^2\mu dx$$

したがって, (6.4)が示された. 一意性は明らか. □

関数ξ_0を$\xi_0=\xi_1/|\xi_1|_w^2$と定めれば, $|\xi_0|_w=1/|\xi_1|_w$ だから, 定理6.1より,

$$(6.7)\qquad L'(\theta)^*[wL'(\theta)\xi_0]+\sigma^2\mu\xi_0 = |\xi_0|_W^2\lambda\sigma^2\mu$$

を得る.

定理 6.2. ξ_0は最適化問題(P3)の一意解である. すなわち, $\xi_0\in C(w,\theta)$であって, $|\xi_0|_w=\min\{|\xi|_w;\xi\in C(w,\theta)\}$.

証明. (6.6)から$\xi_0\in C(w,\theta)$であることはすぐわかる. $\xi\in\mathcal{D}$に対して,

$$(6.8)\qquad 0\leqq|\xi-\xi_0|_W^2$$
$$=|\xi|_W^2+|\xi_0|_W^2-2\left\{\int\xi\xi_0\sigma^2\mu dx+\int(L'(\theta)\xi)(L'(\theta)\xi_0)wdx\right\}$$
$$=|\xi|_W^2+|\xi_0|_W^2-2\left\{\int\xi\xi_0\sigma^2\mu dx+\int\xi L'(\theta)^*[wL'(\theta)\xi_0)]dx\right\}$$
$$=|\xi|_W^2+|\xi_0|_W^2-2|\xi_0|_W^2\int\xi\lambda\sigma^2\mu dx.$$

$\xi\in C(w,\theta)$に対して$\mathcal{D}$の中に列$(\phi(n))$で$|\xi-\phi(n)|_w\to0$, したがって, $|\phi(n)|_w\to|\xi|_w$, $\int\phi(n)\lambda\sigma^2\mu dx\to1$となるものが存在する. 近似の議論によって, (6.8)から, $\xi\in C(w,\theta)$に対して,

$$0 \leqq |\xi|_w^2 + |\xi_0|_w^2 - 2|\xi_0|_w^2 = |\xi|_w^2 - |\xi_0|_w^2$$

となることがわかる．さらに，一意性は明らかである．□

最後に，例をあげる．セクション5であげたOrnstein-Uhlenbeck の拡散過程を考える．関数$w(x)$にたいして最適解を与える方程式(6.4)は

(6.9) $\quad -\xi'' - (w'/w)\xi' + [4\theta^2 x^2 + 2(w'/w)\theta x + 2\theta + 4\mu/w]\xi = 4\mu x/w$

ここで，"′"はxに関する微分を表し，$\mu = (\theta/\pi)^{1/2}\exp(-\theta x^2)$.方程式(6.9)の数値解を得るためには変数$\xi$を適当に変換するのが良いがここでは詳細には立ち入らない．$\theta=1/2$のときMLEの漸近分散は1であるが，いろいろな$w(x)$に対する最適w-ロバストM-推定量の漸近分散を示すと下の表のようになる．この表では速く減少する$w(x)$に対応する最適w-ロバストM-推定量の漸近分散は小さいことがわかる．

表2. 関数$w(x)$とそれに対応する最適w-ロバストM-推定量の漸近分散

$w(x)$	漸近分散
1	1.48
$1/2(1+\|x\|^2)$	1.13
$1/2(1+\|x\|^4)$	1.03
$1/2(1+\|x\|^6)$	1.02
$\exp(-x^2/2)$	1.00

有益な助言を頂いた稲垣宣生教授，近藤正男先生，谷口正信先生に感謝の意を表したい．また，大阪大学の吉田稔先生にはいろいろな点で相談にのって頂き，林利治氏には計算機プログラムに関して助言を頂いたことにお礼申し上げたい．

参考文献.

[1] Bustos, O. (1982). General M-estimates for contaminated pth-order autoregressive processes: consistency and asymptotic nomality. Z. Wahrsch. 59, 491-504.

[2] Bustos, O. and Yohai, V.J. (1986). Robust estimates for ARMA models. J. Amer. Statist. Assoc. 81, 155-168.

[3] Denby, L. and Martin, R.D. (1979). Robust estimation of the first order autoregressive parameter. J. Amer. Statist. Assoc. 74, 140-146.

[4] Feigin, P.D. (1985). Stable convergence of semimartingales, Stoch. Processes Appl., 19, 125-134.

[5] Gihman, I.I. and Skorohod, A.V. (1972). Stochastic differential equations, Springer-Verlag.

[6] Hampel, F.R. (1974). The influence curve and its role in robust estimation. J. Amer. Statist. Assoc. 69, 383-393.

[7] Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J. and Stahel, W.A., (1986). Robust statistics, Wiley.

[8] Ikeda, N. and Watanabe, S. (1981). Stochastic differential equations anddiffusion processes, North-Holland.

[9] Krylov, N.V., (1980). Controlled diffusion processes, Springer-Verlag.

[10] Künsch, H. (1984). Infinitesimal robustness for autoregressive processes, Ann. Statist. 12, 843-863.

[11] Kutoyants, Yu.A. (1977). Estimation of the drift coefficient parameter of a diffusion in the smooth case, Theor. Prob. Appl., 22, 399-406.

[12] Kutoyants, Yu.A. (1978). Estimation of a parameter of a

diffusion process, Theor. Probab. Appl., 641-649.

[13] Kutoyants, Yu.A. (1984). Parameter estimation for stochastic processes, Translated and edited by B.L.S.Prakasa Rao, Heldermann Verlag Berlin.

[14] Lánska, V. (1979). Minimum contrast estimation in diffusion processes, J. Appl. Prob. 16, 65-75.

[15] Lepingle, D. (1978). Sur le comportement asymptotique des martingales locales, Lect. Notes in Math., 649, 148-161, Springer-Verlag.

[16] Liptser, R.Sh. and Shiryayev, A.N. (1977). Statistics of random processes, Vol. I, Springer-Verlag.

[17] Martin, R.D. and Yohai, V.J. (1985). Robustness in time series and estimating ARMA models. Handbook of Statistics, Vol. 5, 119-155.

[18] McKeague, I.W. (1984). Estimation for diffusion processes under misspecified models, J. Appl. Prob., 21, 511-520.

[19] Powell, M.J.D., ed. (1982). Nonlinear optimization 1981, Academic Press.

[20] Rao, B.L.S.P. and Rubin, H., (1981). Asymptotic theory of estimation in non-linear stochastic differential equations, Sankhya, Vol. 43, Ser. A, Pt. 2, 170-189.