数理解析研究所講究録
1039巻 1998年 141-151

情報解析における基本不等式

梅垣壽春
Hisaharu UMEGAKI
東京工業大学名誉教授
中国東北師範大学客座教授

§1. C.E.Shannonが1948年に大論文[3]を発表し，情報理論という新領域を世に誕生させたことは周知のことであるが，この論文によって20世紀に於ける情報時代がstartしたことは万人の知る所であり，1948年が情報元年と呼ばれる所以である。更に，大書す可きは，このnew fieldsのstarting pointであり，基盤であるのがEntropy概念である。このことについて詳論することは重要であり，興味あることであるが，既に多数の書物で論じられているので，本論では略する。

そこでは，Entropy定式を構成するための数学，特に函数解析における基本的不等式が如何に有効であるかの一端を論じ度い。

§2. 基本不等式とK-L情報量

Entropy を論ずる際に基本的役割を演ずる三つの重要な不等式を述べよう。

1°. Schwarz 不等式

$$\left| \int_\Omega f(\omega) g(\omega) d\mu(\omega) \right|^2 \leqq \int_\Omega |f(\omega)|^2 d\mu(\omega) \cdot \int_\Omega |g(\omega)|^2 d\mu(\omega)$$

$$\text{for } \forall f, g \in L^2(\Omega) \qquad (2.1)$$

茲で，$\Omega = (\Omega, \mathcal{L}_\Omega, \mu)$ は σ-finite 測度空間，特に

等号 (for $f \neq 0 \neq g$) の成立

$\iff \exists \lambda$ (スカラー)； $f = \lambda g$ (a.e.)．

2°. 対数不等式

$$\log t \geqq 1 - t^{-1} \quad (t > 0) \qquad (2.2)$$

等号 $\iff t = 1$

3°. Kullback-Leibler (K-L) 不等式

$$\int_\Omega f(\omega) \log g(\omega) d\mu(\omega) \leqq \int_\Omega f(\omega) \log f(\omega) d\mu(\omega) \qquad (2.3)$$

for $\forall f, g \in L^1(\Omega)^+$, $\|f\|_1 = \|g\|_1$, $\mathrm{supp} f \subset \mathrm{supp}\, g$,

$$\text{等号の成立} \Longleftrightarrow f = g \quad (a.e.)$$

茲で，Ω は $1°$ と同じ．

$1°$ は well-known，$2°$ は容易，$3°$ は $2°$ から容易。

この K-L 不等式から，$\forall$(可算)確率分布

$$\mathbb{p} = (p_1, p_2, \cdots),\quad \mathbb{q} = (q_1, q_2, \cdots)$$

に対して

$$S(\mathbb{p}/\mathbb{q}) \triangleq \sum (p_j \log p_j - p_j \log q_j) \geqq 0 \quad (2.4)$$

且つ，　等号の成立 $\Longleftrightarrow p_j = q_j \; (j = 1, 2, \cdots)$

が云える。

式 (2.4) の値 $S(\mathbb{p}/\mathbb{q})$ を $\mathbb{q}$ に関する $\mathbb{p}$ の相対エントロピー の名称で呼ばれる。これは Kullback と Leibler が Shannon の 5年後の 1953年に導入した量で Kullback-Leibler (K-L) 情報量と呼ばれる。

§3. GKY相対エントロピー $S(\mu/\nu)$。

可測空間 $(\Omega, \mathcal{L}_\Omega)$ 上の確率測度の対 μ, ν に対して，$\mathcal{L}$ の σ-部分集合体 $\mathcal{F}$ を与へるとき

$$S_{\mathcal{F}}(\mu/\nu) \doteq \sup\left\{ \sum_{A \in \mathcal{A}} \mu(A) \log \frac{\mu(A)}{\nu(A)} ; \mathcal{A} \in \mathcal{P}_{\mathcal{F}}(\Omega) \right\}$$

を対 μ, ν の $\mathcal{F}$ に関するGKY-相対エントロピーという。ここで $\mathcal{P}_{\mathcal{F}}(\Omega)$ は $\mathcal{F}$-可測な有限分割 $\mathcal{A}$ の全体とする。ここで，分数 $a/0 = \infty\ (a>0)$ とする。$\mathcal{F} = \mathcal{L}$ のときは単に

$$S(\mu/\mu) = S_{\mathcal{L}}(\mu/\nu)$$

とかき，対 μ, ν のGKY相対エントロピーという。「GKY」は，この概念の発見者 Gelfand, Kolmogorov と Yaglom の頭文字である。これに関する基本定理を述べると，

定理 3.1. 可測空間 $(\Omega, \mathcal{L})$ 上の確率測度の対 μ, ν に対して，次の(1),(2)が成立：

(1) $\mu \ll \nu \Rightarrow$

$$S(\mu/\nu) = \int \left(\frac{d\mu}{d\nu} \log \frac{d\mu}{d\nu} \right) d\mu = \int \log \frac{d\mu}{d\nu}\, d\mu$$

(2) $\mu \not\ll \nu \Rightarrow S(\mu/\nu) = \infty$.

これの証明はKalianpurを改良した塚田真氏の方法で与へてある cf.[7]。

§4. 作用素間の相対エントロピー

Quantum Mechanical Statistical Operators A, B が可分Hilbert空間 H 上に act しているとする。これはSchatten記号によって表される:

$$A = \sum a_n \varphi_n \otimes \overline{\varphi_n}, \quad B = \sum b_n \psi_n \otimes \overline{\psi_n} \qquad (4.1)$$

茲で, $\sum a_n = \sum b_n = 1$, $a_n, b_n \downarrow 0$ 且つ $\{\varphi_n\}$, $\{\psi_n\}$ は共にCONS(完全正規直交系)と設定しておく。このとき

A のエントロピー: $S(A) = -\sum a_n \log a_n$,

A の B に関する相対エントロピー は

$$S(A/B) = \sum (a_n \log a_n - a_n \log b_n) \qquad (4.2)$$

によって与えられる。これらはTraceを用いれば

$$S(A) = -\mathrm{Tr}(A \log A), \quad S(A/B) = \mathrm{Tr}(A \log A - A \log B) \qquad (4.3)$$

と表される。(以上，1962年 Umegaki [6])。

註．最近 Tr を用いず，抽象的・一般的な議論が T. Furuta, M. Fujii, J. Fujii 等によって為されている。

茲で，von Neumann-Schatten 記号 $\varphi_i \otimes \overline{\psi_j}$ 等を用いて，

$$A \log B = \left(\sum_i a_i \varphi_i \otimes \overline{\varphi_i}\right)\left(\sum_j \log b_j \psi_j \otimes \overline{\psi_j}\right)$$

$$= \sum_{i,j} a_i \log b_j (\varphi_i \otimes \overline{\varphi_i})(\psi_j \otimes \overline{\psi_j})$$

$$= \sum_{i,j} (a_i \log b_j) \langle \varphi_i, \psi_j \rangle \varphi_i \otimes \overline{\psi_j},$$

となり，Trace を用いて

$$Tr(A \log B) = \sum_{i,j} a_i \log b_j \langle \varphi_i, \psi_j \rangle \cdot Tr(\varphi_i \otimes \overline{\psi_j})$$

$$= \sum_{i,j} a_i \log b_j |\langle \varphi_i, \psi_j \rangle|^2$$

これらに対して，函数 log. の凹性を用いれば

$$Tr(A \log A - A \log B)$$

$$= \sum_i a_i \log a_i - \sum_i a_i \left(\sum_j |\langle \varphi_i, \psi_j \rangle|^2 \log b_j\right)$$

$$\geqq \sum_i a_i \log a_i - \sum_i a_i \log \langle \varphi_i, \sum_j b_j (\psi_j \otimes \overline{\psi_j}) \varphi_i \rangle$$

$$= \sum_i a_i \log a_i - \sum_i a_i \log \langle \varphi_i, B\varphi_i \rangle$$

$$= \sum \langle \varphi_i, B\varphi_i \rangle \left(\frac{a_i}{\langle \varphi_i, B\varphi_i \rangle} \log \frac{a_i}{\langle \varphi_i, B\varphi_i \rangle} \right) \quad (4.4)$$

茲で対数不等式(2.2)を用いて

$$\geqq \sum \langle \varphi_i, B\varphi_i \rangle \left(-1 + \frac{a_i}{\langle \varphi_i, B\varphi_i \rangle} \right)$$

$$= \sum - \langle \varphi_n, B\varphi_n \rangle + \sum a_n = 0$$

故に

$$S(A/B) = \mathrm{Tr}(A \log A - A \log B) \geqq 0 \qquad (4.5)$$

$S(A/B) = 0$ のとき(4.4)の値 $= 0$。従って

$$\langle \varphi_n, A\varphi_n \rangle = \langle \varphi_n, B\varphi_n \rangle, \ n = 1, 2, \cdots \quad (4.6)$$

このとき Umegaki Conditional Expectation によって

$$B^e \triangleq E\left[B / \left(\{ \varphi_n \otimes \overline{\varphi_n} \}_{n=1}^{\infty} \right)' \right]$$

とおく, cf. Umegaki[4], 及び解説[1]参照。

ここで $\left(\{ \varphi_n \otimes \overline{\varphi_n} \}_{n=1}^{\infty} \right)'$ は max abel in $\mathcal{B}(H)$ であるから

$$\left(\{ \varphi_n \otimes \overline{\varphi_n} \}_{n=1}^{\infty} \right)' = \left(\{ \varphi_n \otimes \overline{\varphi_n} \}_{n=1}^{\infty} \right)'', \triangleq \mathfrak{A}$$

とおくと

$$B^e = E[B/\mathfrak{A}]$$

これらの記号を用いて

$$\langle \varphi_n, B\varphi_n \rangle = \mathrm{Tr}((B\varphi_n) \otimes \overline{\varphi_n}) = \mathrm{Tr}(B(\varphi_n \otimes \overline{\varphi_n}))$$

$$= \mathrm{Tr}(B^e(\varphi_n \otimes \overline{\varphi_n})) = \langle \varphi_n, B^e \varphi_n \rangle.$$

これと等式(4.6)を用いて，次の結論に到着する。

定理 4.1. 量子統計的状態を表す二つの自己共軛作用素 A, B が共に縮退のないものとし，これの相対エントロピー

$$S(A/B) = \mathrm{Tr}(A \log A - A \log B)$$

は $\geqq 0$ であり，特に $S(A/B) = 0$ ならば $A = E[B/\mathfrak{a}]$ となる。

上記定理の構成の途上に於いて，

$$p_n = \varphi_n \otimes \overline{\varphi_n} \tag{4.7}$$

は 1-dim proj. で $p_i \perp p_j$ $(i \neq j)$ である。このことを用いて

$$\begin{aligned} &B_0 \triangleq B,\ B_1 \triangleq p_1 B_0 p_1 + (1-p_1) B_1 (1-p_1) \\ &B_n \triangleq p_n B_{n-1} p_n + (1-p_n) B_{n-1} (1-p_n),\ n = 0, 1, 2, \cdots \end{aligned} \tag{4.8}$$

とおくと，確率過程に於けるmartingaleと，その収束定理の非可換判の一つが次の様にformulateされる。

定理 4.2. (Martingale収束).

$\{\varphi_n\}$をCONS, Bを量子統計的状態を表す自己共役作用素とし, p_nとB_nを(4.7)と(4.8)によって定めると

$$B_n = E[B_{n-1}/\{p_n, 1-p_n\}']$$
$$= E[B/\{p_1, p_2, \cdots, n\}'], \quad n=1,2,\cdots$$

となり, $\{B_n\}$はoperator-martingaleとなり

$$\lim_{n\to\infty} B_n = \lim_{n\to\infty} E[B/\{p_1, p_2, \cdots, p_n\}''],$$

茲でlimは作用素強収束。

この定理によって, A, Bを定理3.1の作用素対とするとき, 次定理が導かれる。

定理 4.3.

$$A = \lim_{n\to\infty} E[B/\{p_1, p_2, \cdots p_n\}'']$$

[註] 上記operator-martingaleに関する議はUmegaki[5]を参照されたい。概念の基は半然確率過程(Doob)であり, 構成baseはvon Neumann理論[2]である。

§5. Schwarz不等式の典型的な

効果として論じることが出来るのは、不確定性関係を表わす、今世紀最大ともいうべき不等式

$$\mathrm{var}(\rho, A)\cdot \mathrm{var}(\rho, B) \geqq (4\pi)^{-2}$$

である，(A, Bは正準交換関係を満たすobservablesの対)．この重大な関係式はSchwarzの不等式を用いて，大変明解に示される。而もこの等号の場合の状態函数に関する必要十分条件がSchwarz不等式の等号の場合の，そのまゝの適用として働き，更に，$(A, B) = (P, Q)$の場合に，等号の方程式を構成することによって興味ある函数解析とFourier解析が展開される。ここでは紙面が尽きたので割愛する。

参考文献

[1] L. Accadi, Quantum Stochastic Processes, ENCYCLOPAEDIA of MATHEMATICS, Vol. 10 (1993).

[2] J. von Neumann, On Rings of Operators, III Ann. of Math., 41 (1939), 94-161.

[3] C. E. Shannon, A Mathematical Theory of Communication, Bell. System Tech. Journ. 27 (1948), 374-423, 623-656.

[4] H. Umegaki, Conditional expectation in an operator algebra, I, Tohoku Math. J. 6 (1954), 177-181.

[5] H. Umegaki, Conditional expectation in an operator algebra II, Ibid 8 (1956), 86-100.

[6] H. Umegaki, Conditional expectation in an operator algebra IV, Kodai Math. J. 14 (1962) 59-85.

[7] 梅垣壽春, 情報数理の基礎, サイエンス社, 1994年.