

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と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、既に多数の事物で論いられているので、本論では略する。

ここでは、Entropy 定式を構成するための数学、特に通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

$$\left| \int_{\Omega} f(\omega) g(\omega) d\mu(\omega) \right|^2 \leq \int_{\Omega} |f(\omega)|^2 d\mu(\omega) \int_{\Omega} |g(\omega)|^2 d\mu(\omega) \\ \text{for } \forall f, g \in L^2(\Omega) \quad (2.1)$$

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

$$\iff \exists \lambda (\text{スカラー}): f = \lambda g \text{ (a.e.)}.$$

2°. 対数不等式

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

$$\text{等号} \iff t = 1$$

3°. Kullback-Leibler (K-L) 不等式

$$\int_{\Omega} f(\omega) \log g(\omega) d\mu(\omega) \leq \int_{\Omega} f(\omega) \log f(\omega) d\mu(\omega) \quad (2.3)$$

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

等号の成立 $\iff f = g$ (a.e.)

茲で, Ω は I^0 と同じ.

1^0 は well-known, 2^0 は容易, 3^0 は 2^0 から容易.

2の K-L 不等式から, $\forall$ (可算) 確率分布

$$p = (p_1, p_2, \dots), \quad q = (q_1, q_2, \dots)$$

に対して

$$S(p/q) \triangleq \sum (p_j \log p_j - p_j \log q_j) \geq 0 \quad (2.4)$$

且つ, 等号の成立 $\iff p_j = q_j$ ($j = 1, 2, \dots$)
が 成 立 する。

式 (2.4) の 値 $S(p/q)$ を q に関する 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, $\mathcal{L}_\Omega$ の σ -部分集合体 $\mathcal{F}$ を与えるとき

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

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。 $\mathcal{F} = \mathcal{L}_\Omega$ のときは単に

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

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 $(\Omega, \mathcal{L}_\Omega)$ 上の確率
測度の対 μ, ν に対して, 次の (1), (2) が成立:

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

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

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

これらの証明は Kallianpur を改良した塚田眞氏の方法でなへてある 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:

$$A = \sum a_n \varphi_n \otimes \bar{\varphi}_n, \quad B = \sum b_n \psi_n \otimes \bar{\psi}_n \quad (4.1)$$

茲で, $\sum a_n = \sum b_n = 1, a_n, b_n \geq 0$ 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

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

によって与えられる。これらは Trace を用いれば

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

と表される。(以上, 1962年 Umegaki [6]).

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

茲で, von Neumann-Schatten 記号 $\varphi_i \otimes \bar{\varphi}_j$ 等を用いて,

$$\begin{aligned} A \log B &= \left(\sum_i a_i \varphi_i \otimes \bar{\varphi}_i \right) \left(\sum_j \log b_j \psi_j \otimes \bar{\psi}_j \right) \\ &= \sum_{i,j} a_i \log b_j (\varphi_i \otimes \bar{\varphi}_i) (\psi_j \otimes \bar{\psi}_j) \\ &= \sum_{i,j} (a_i \log b_j) \langle \varphi_i, \psi_j \rangle \varphi_i \otimes \bar{\psi}_j, \end{aligned}$$

となり, Trace を用いて

$$\begin{aligned} \text{Tr}(A \log B) &= \sum_{i,j} a_i \log b_j \langle \varphi_i, \psi_j \rangle \cdot \text{Tr}(\varphi_i \otimes \bar{\psi}_j) \\ &= \sum_{i,j} a_i \log b_j |\langle \varphi_i, \psi_j \rangle|^2 \end{aligned}$$

これらに対して, 函数 $\log$ の凹性を用いれば

$$\begin{aligned} &\text{Tr}(A \log A - A \log B) \\ &= \sum_i a_i \log a_i - \sum_i a_i \left(\sum_j |\langle \varphi_i, \psi_j \rangle|^2 \log b_j \right) \\ &\geq \sum_i a_i \log a_i - \sum_i a_i \log \left\langle \varphi_i, \sum_j b_j (\psi_j \otimes \bar{\psi}_j) \varphi_i \right\rangle \\ &= \sum_i a_i \log a_i - \sum_i a_i \log \langle \varphi_i, B \varphi_i \rangle \\ &= \sum \langle \varphi_i, B \varphi_i \rangle \left(\frac{a_i}{\langle \varphi_i, B \varphi_i \rangle} \log \frac{a_i}{\langle \varphi_i, B \varphi_i \rangle} \right) \quad (4.4) \end{aligned}$$

ここで対数不等式 (2.2) を用いて

$$\begin{aligned} &\geq \sum \langle \varphi_i, B \varphi_i \rangle \left(-1 + \frac{a_i}{\langle \varphi_i, B \varphi_i \rangle} \right) \\ &= \sum -\langle \varphi_n, B \varphi_n \rangle + \sum a_n = 0 \end{aligned}$$

故に

$$S(A/B) = \text{Tr}(A \log A - A \log B) \geq 0 \quad (4.5)$$

$S(A/B) = 0$ のとき (4.4) の値 $= 0$ 。従って

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

このとき Umegaki Conditional Expectation によって

$$B^e \triangleq E[B / (\{\varphi_n \otimes \bar{\varphi}_n\}_{n=1}^{\infty})']$$

とおく, cf. Umegaki[4], 及び解説[1]参照。

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

$$(\{\varphi_n \otimes \bar{\varphi}_n\}_{n=1}^{\infty})' = (\{\varphi_n \otimes \bar{\varphi}_n\}_{n=1}^{\infty})'' \triangleq \mathcal{A}$$

とおくと

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

これらの記号を用いて

$$\langle \varphi_n, B \varphi_n \rangle = \text{Tr}((B \varphi_n) \otimes \bar{\varphi}_n) = \text{Tr}(B(\varphi_n \otimes \bar{\varphi}_n))$$

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

これと等式 (4.6) を用いて, 次の結論に到着する.

定理 4.1. 量子統計的収束を表す二つの自己共轭作用素 A, B が共に縮退のないものとし, 2つの相対エントロピー

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

は ≥ 0 であり, 特に $S(A/B) = 0$ ならば " $A = E[B/\mathcal{Q}]$ " となる.

上記定理の構成の途上において,

$$p_n = \varphi_n \otimes \overline{\varphi}_n \quad (4.7)$$

は 1-dim proj. で $p_i \perp p_j$ ($i \neq j$) である.

このことを用いて

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

とみると, 確率過程に於ける martingale と, その収束定理の非可換判の 하나가 次の様に formulate される.

定理 4.2. (Martingale 収束).

$\{p_n\}$ を CONS, B を量子統計的状态を表す
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

$$B_n = E[B_{n-1} / \{p_n, 1-p_n\}']$$

$$= E[B / \{p_1, p_2, \dots, p_n\}'], \quad n=1, 2, \dots$$

となり, $\{B_n\}$ は operator-martingale となり

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

ここで $\lim$ は作用素-強収束.

この定理において, A, B を定理 3.1 の作用素
対とすると, 次定理が導かれる.

定理 4.3.

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

註 上記 operator-martingale に関する議論
は Umegaki [5] を参照されたい. 概念の基は半
強確率過程 (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、今世紀最大とも云う可き不等式

$$\text{var}(p, A) \cdot \text{var}(p, B) \geq (4\pi)^{-2}$$

である、(A, B は正準交換関係を満たす observables
 の対)。この重要な関係式は Schwarz の不等式
 を用いて、大変明解に示めされる。而もこの
 等号の場合の状態函数に因する必要十分条
 件が Schwarz 不等式の等号の場合の、そのままの
 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。ここでは紙面が盡きたので
 割愛する。

参考文献

- [1] L. Accardi, 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.