一般社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
IEICE Technical Report
IT2013-52 (2014-1)

乱数変換に対する二次漸近解析

熊谷　亘[†]　林　正人[†,††]

† 名古屋大学 多元数理科学研究科
〒 464-8602 愛知県名古屋市千種区不老町
†† シンガポール国立大学
3 Science Drive 2, Singapore 117543, Singapore
E-mail: †wkumagai1001@gmail.com, ††masahito@math.nagoya-u.ac.jp

あ a @　本発表では確率分布 P から生成されるの独立同一乱数を、Q から生成される独立同一乱数へ変換する際の最適生成レートに関して論じる。確率分布 Q または P が一様分布である場合は、この問題はそれぞれ intrinsic randomness および resolvability と呼ばれており、一次だけでなく二次の漸近解析もなされてきた。一方、P および Q の双方が非一様分布である場合、その二次漸近論は扱われてこなかった。本発表では P および Q が有限集合上の一様とは限らない確率分布とし、それらの間の最適変換レートを行う。

キーワード　乱数変換、二次漸近論、漸近展開

Wataru KUMAGAI[†] and Masahito HAYASHI[†,††]

† Nagoya University
464-8602, Furocho, Chikusaku, Nagoya, Japan
†† National University of Singapore
3 Science Drive 2, Singapore 117543, Singapore
E-mail: †wkumagai1001@gmail.com, ††masahito@math.nagoya-u.ac.jp

Abstract　We treat a random number generation from an i.i.d. probability distribution of P to that of Q. When Q or P is a uniform distribution, the problems have been well-known as the uniform random number generation and the resolvability problem respectively, and analyzed not only in the context of the first order asymptotic theory but also that in the second asymptotic theory. On the other hand, when both P and Q are not a uniform distribution, the second order asymptotics has not been treated. In this paper, we focus on the second order asymptotics of a random number generation for arbitrary probability distributions P and Q on a finite set. In particular, we derive the optimal second order generation rate under an arbitrary permissible confidence coefficient.

Key words　Random number conversion, second-order asymptotic theory, asymptotic expansion

1. 導　　入

乱数変換は情報理論の中で最も基本的な問題の一つである。乱数変換では、与えられた確率分布 P_n を対象となる確率分布 Q_n に所定の精度で変換することを目的としている。確率分布 Q_n または P_n が一様分布である場合は、この問題はそれぞれ intrinsic randomness および resolvability と呼ばれており、精密な評価がなされてきた。例として、U_2 をサイズ 2 の一様分布とするとき、独立同一分布 P^n から一様乱数 $U_2^{\lfloor an \rfloor}$ を生成するとき、誤差が 0 に向かうという条件の下での最適レート a は P のエントロピー $H(P)$ で与えられる。乱数変換の問題は P_n もしくは Q_n の一方が一様分布である場合には、一次だけでなく二次の漸近理論の文脈で解析がなされており、一般に一次および二次の最適変換レートは情報スペクトルで記述される [1], [3], [4]. 他方、P_n および Q_n の双方が非一様分布である場合、最適変換レートに関する研究は十分になされてこなかった。本発表では P_n または Q_n を有限集合上の一様分布とは限らない独立同一分布であるとし、それら間の最適変換レートを精度制約の下で明示的に導出する。本発表では、二次の変換レートを導出するために、二つの確率分布の間の半順序である majorization の考え方を用いる。majorization は通常の乱数の決定的変換を含むより一般的な概念であり、決定的変換に関する逆定理を証明する際に重要な役割を果たす。majorization は従来の情報理論の枠組みでは現れないが、量子情報理論にお

ける量子エンタングル状態の LOCC 変換問題として自然に表れることが知られている。

2. 問題の定式化

本節ではいくつかの記法を導入し、問題を正確に定式化する。有限集合 $\mathcal{X}$ 上の確率分布 P と写像 $W:\mathcal{X}\to\mathcal{Y}$ に対し、$\mathcal{Y}$ 上の確率分布が $W(P)(y):=\sum_{x\in W^{-1}(y)}P(x)$ によって定まる。この写像から定まる確率分布の変換を以下では決定的変換と呼ぶことにする。

ここで、同一の有限集合 $\mathcal{Y}$ 上の確率分布の間の忠実度（もしくは Bhattacharyya 係数）と呼ばれる量を以下の様に導入する。

$$F(Q,Q'):=\sum_{y\in\mathcal{Y}}\sqrt{Q(y)Q'(y)}. \tag{1}$$

このとき F は二つの確率分布がどれ程近いかを表しており、Hellinger 距離 d_H は $d_H(\cdot,\cdot)=\sqrt{1-F(\cdot,\cdot)}$ の様に表される. 我々の目的は以下の最大変換可能数を解析することである。

$$L(P,Q|\nu):=\max\{L|F(W(P),Q^L)\geqq\nu, W:\mathcal{X}\to\mathcal{Y}^L\} \tag{2}$$

この量は精度制約の下で確率分布 P から最大で Q の独立同一分布がどの程度生成できるかを表す。$\mathcal{X}$ 上の確率分布 P から $\mathcal{Y}$ 上の確率分布 Q への最大忠実度を

$$F(P\to Q):=\max\{F(W(P),Q)|W:\mathcal{X}\to\mathcal{Y}\}, \tag{3}$$

によって定義するとき、最大変換可能数は以下の様に書くことができる。

$$L(P,Q|\nu)=\max\{L|F(P\to Q^L)\geqq\nu\}. \tag{4}$$

次に majorization の概念を導入する。有限集合上の確率分布 P に対し、$P^{\downarrow}$ は P の成分を降順に並べた $\{1,...,|\mathcal{X}|\}$ 上の確率分布 $\{P_i^{\downarrow}\}_{i=1}^{|\mathcal{X}|}$ とする。二つの確率分布 P と Q が任意の l に対し $\sum_{i=1}^{l}P_i^{\downarrow}\leqq\sum_{i=1}^{l}Q_i^{\downarrow}$ をみたすとき、"P は Q に majorize される" といい $P\prec Q$ と表す。ここで、必要なら $P^{\downarrow}$ または $P^{\downarrow}$ に余分に 0 成分を加えることで、P と Q は同一の集合上に定義されていなくとも二つの確率分布の間に majorization の半順序関係を定義することができる [2]。majorization 条件の下での有限集合上の確率分布 P から Q への最大忠実度を

$$F^M(P\to Q):=\max\{F(P',Q)|P\prec P'\ on\ \mathcal{Y}\} \tag{5}$$

と定める。決定的変換 $W:\mathcal{X}\to\mathcal{Y}$ に対し $P\prec W(P)$ が成り立つので,

$$F^M(P\to Q)\geqq F(P\to Q) \tag{6}$$

を得る。多くの場合において $F^M(P\to Q)$ は $F(P\to Q)$ より扱いやすい。majorization の概念は量子情報理論において自然に表れるが、ここでは立ち入らない [7], [8]。ただ、乱数変換の問題は量子情報理論における文脈 (LOCC 変換の問題) に翻訳することができるを指摘しておく [4], [6]。

3. 一様分布に対する二次漸近解析

本節以降では P^n から $Q^{an+b\sqrt{n}}$ への変換に関する漸近解析を行う。特に最大変換可能数 $L_n(P,Q|\nu)$ の二次漸近展開公式を導出する。本節では P もしくは Q がサイズ 2 の一様分布 U_2 である場合を扱う。

始めに $P=U_2$ とし resolvability 問題を扱う。任意の精度係数 $0<\nu<1$ にたいし、以下の二次漸近展開が得られる。

[Theorem 1] Q を有限集合上の非一様確率分布とする。そのとき以下の展開が成り立つ。

$$\begin{aligned}&L_n(U_2,Q|\nu)\\&=H(Q)^{-1}n-\sqrt{\frac{V(Q)}{H(Q)^3}}\Phi^{-1}(\nu^2)\sqrt{n}+o(\sqrt{n}),\end{aligned} \tag{7}$$

ここで Φ は標準正規累積分布関数であり、$H(Q)$ は Q のエントロピーとし、

$$V(Q):=\sum_{x\in\mathcal{X}}Q(x)(-\log Q(x)-H(P))^2 \tag{8}$$

とした。

上記の定理を得るために以下の命題が重要である。

[Proposition 2] Q を有限集合上の非一様確率分布とする。そのとき以下が成り立つ。

$$\begin{aligned}&\lim_{n\to\infty}F(U_2^n\to Q^{H(Q)^{-1}n+b\sqrt{n}})\\&=\lim_{n\to\infty}F^M(U_2^n\to Q^{H(Q)^{-1}n+b\sqrt{n}})\\&=\sqrt{\Phi\left(\frac{-H(Q)^{\frac{3}{2}}b}{\sqrt{V(Q)}}\right)}.\end{aligned} \tag{9}$$

式 (7) における二次の漸近レートは上の最大忠実度の極限値に関する命題と簡単な計算により導かれる。以後現れる二次の漸近展開公式も同様に、最大忠実度の極限値の解析によって得られる。命題 2 より、$a=H(P)/H(Q)$ のとき最大忠実度の極限値は二次の変換レート b に依存することがわかる。一方、$a>H(P)/H(Q)$ または $a<H(P)/H(Q)$ のときは、二次の変換レート b に依存せず、0 または 1 になる。

次に $Q=U_2$ とし intrinsic randomness 問題を扱う。任意の精度係数 $0<\nu<1$ にたいし、以下の二次漸近展開が得られる。

[Theorem 3] P を有限集合上の非一様確率分布とする。そのとき以下の展開が成り立つ。

$$\begin{aligned}&L_n(P,U_2|\nu)\\&=H(P)n-\sqrt{V(P)}\Phi^{-1}(\nu^2)\sqrt{n}+o(\sqrt{n}).\end{aligned} \tag{10}$$

4. 非一様分布に対する二次漸近解析

本節では P と Q が共に非一様分布である場合に、最大変換可能数 $L_n(P,Q|\nu)$ の二次漸近展開公式を導出する。本節での結果は $V(P)$ と $V(Q)$ の両方が 0 でないことを用いるので、そ

のために P と Q が共に非一様分布であることは本質的である。従って本節の結果は前節の一様分布に対する結果を含むわけではないことに注意する。

まず、非一様分布 P,Q に対しいくつかの記法を導入する。

$$N_{P,Q,b} := N\left(D_{P,Q}b, C_{P,Q}\right), \tag{11}$$

$$\Phi_{P,Q,b}(x) := \Phi\left(\sqrt{C_{P,Q}}^{-1}(x - D_{P,Q}b)\right), \tag{12}$$

$$\begin{aligned} I_{P,Q,b}(x) &:= \int_{-\infty}^{x} \sqrt{N(t)N_{P,Q,x}(t)}dt \\ &= \sqrt{\frac{2\sqrt{C_{P,Q}}}{1+C_{P,Q}}} e^{-\frac{(D_{P,Q}b)^2}{4(1+C_{P,Q})}} \\ &\quad \times \Phi\left(\sqrt{\frac{1+C_{P,Q}}{2C_{P,Q}}}\left(x - \frac{D_{P,Q}b}{1+C_{P,Q}}\right)\right), \end{aligned} \tag{13}$$

$$\begin{aligned} I_{P,Q,b}(\infty) &:= \int_{-\infty}^{\infty} \sqrt{N(t)N_{P,Q,x}(t)}dt \\ &= \sqrt{\frac{2\sqrt{C_{P,Q}}}{1+C_{P,Q}}} e^{-\frac{(D_{P,Q}b)^2}{4(1+C_{P,Q})}} \end{aligned} \tag{14}$$

ただし $N(\mu,v)$ は平均 μ、分散 v の正規分布であり、$N := N(0,1)$、$C_{P,Q} := \frac{H(P)}{V(P)}\left(\frac{H(Q)}{V(Q)}\right)^{-1}$、$D_{P,Q} := \frac{H(Q)}{\sqrt{V(P)}}$ と定めた。ここで $\Phi_{P,Q,b}$ は $N_{P,Q,b}$ の累積分布関数である。

以下で最大変換可能数 $L_n(P,Q|\nu)$ の二次漸近展開を導出するために、問題を $C_{P,Q}$ の値に応じて三つに分割する。以下が第一の場合である。

[Theorem 4] $C_{P,Q} > 1$ であるとき、方程式

$$\frac{N(x)}{N_{P,Q,b}(x)} = \frac{\Phi(x)}{\Phi_{P,Q,b}(x)} \tag{15}$$

は x に関して唯一解 $\alpha_b \in \mathbb{R}$ を持つ。関数 $\mathrm{r}F_1 : \mathbb{R} \to [0,1]$ が

$$F_1(b) = \sqrt{\Phi(\alpha_b)\Phi_{P,Q,b}(\alpha_b)} + I_{P,Q,b}(\infty) - I_{P,Q,b}(\alpha_b), \tag{16}$$

で定義されるとき、以下の漸近展開が成り立つ。

$$L_n(P,Q|\nu) = (H(P)/H(Q))n + F_1^{-1}(\nu)\sqrt{n} + o(\sqrt{n}). \tag{17}$$

定理 4 は定理 1 と同様に、以下の最大忠実度の極限値に関する解析より得られる。

[Proposition 5] $C_{P,Q} > 1$ のとき以下が成り立つ。

$$\begin{aligned} &\lim_{n\to\infty} F(P^n \to Q^{\frac{H(P)}{H(Q)}n+b\sqrt{n}}) \\ &= \lim_{n\to\infty} F^M(P^n \to Q^{\frac{H(P)}{H(Q)}n+b\sqrt{n}}) = F_1(b). \end{aligned} \tag{18}$$

また、$C_{P,Q} < 1$ および $C_{P,Q} = 1$ の場合も以下の様に二次の漸近展開を得る。

[Theorem 6] $C_{P,Q} < 1$ のとき

$$\frac{N_P(x)}{N_{P,Q,b}(x)} = \frac{1-\Phi(x)}{1-\Phi_{P,Q,b}(x)} \tag{19}$$

は x に関する唯一解 $\beta_b \in \mathbb{R}$ を持つ。関数 $F_2 : \mathbb{R} \to [0,1]$ が

$$F_2(b) = I_{P,Q,b}(\beta_b) + \sqrt{(1-\Phi(\beta_b))(1-\Phi_{P,Q,b}(\beta_b))}, \tag{20}$$

で定義されるとき、以下の漸近展開が成り立つ。

$$L_n(P,Q|\nu) = (H(P)/H(Q))n + F_2^{-1}(\nu)\sqrt{n} + o(\sqrt{n}) \tag{21}$$

[Theorem 7] $C_{P,Q} = 1$ のとき、以下の漸近展開が成り立つ。

$$L_n(P,Q|\nu) = \frac{H(P)}{H(Q)}n + \sqrt{\frac{8V(P)\log\nu^{-1}}{H(Q)}}\sqrt{n} + o(\sqrt{n}). \tag{22}$$

本節の最初に、本節の結果は前節の結果を含んではいないと述べたが、$Q \to U_2$ の極限をとることにより、定理 4 の漸近展開の二次のレートは定理 1 のそれに収束することを示すことができる。同様に $P \to U_2$ なる極限をとることで、定理 6 の漸近展開の二次のレートは定理 3 のそれに収束することを示すことができる。

5. 結論と課題

本論文では乱数変換に関する二次漸近解析を行ってきた。既存研究においては、変換前または変換後の確率分布が一様分布であることが仮定されていた [3], [5]。一方で、ここでは一般の有限集合上の確率分布を扱った。特に P の独立同一分布から Q それへの最大変換可能数に対しで、二次の漸近展開を与えた。変換前の確率分布 P^n の n が大きいとき、最大変換可能数は定義からでは計算が非常に困難になる。一方で我々が導出した漸近公式は n が大きいときに精密な評価を与えることができるという点で、有益な結果である。最後にいくつかの今後の方向性に関して述べる。今回取り扱った問題設定では、変換前の分布が独立同一であることを仮定したが、相関のあるようなもっと一般の分布を扱うことが今後の問題として考えられる。また今回は漸近論を扱ったが、実際的に扱える乱数は有限サイズであるので、有限長の領域における解析も考察すべきであると考えられる。

Acknowledgment

WK acknowledges support from Grant-in-Aid for JSPS Fellows No. 233283. MH is partially supported by a MEXT Grant-in-Aid for Scientific Research (A) No. 23246071. The Center for Quantum Technologies is funded by the Singapore Ministry of Education and the National Research Foundation as part of the Research Centres of Excellence programme.

文　献

[1] T. S. Han, *Information-Spectrum Methods in Information Theory*. New York, Springer, 2003.
[2] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*. New York, Academic Press, 1979.
[3] R. Nomura and T. S. Han, IEEE Trans. Inf. Theory, vol. 59, pp. 1-16, Jan. 2013.
[4] M. Hayashi, IEEE Trans. Inf. Theory, vol. 54, pp. 4619-4637, Oct. 2008.
[5] M. Hayashi, IEEE Trans. Inf. Theory, vol. 52, pp. 1904-1921, May 2006.
[6] C. H. Bennett *et al.*, Phys. Rev. A, vol. 53, pp. 2046-2052, 1996.
[7] G. Vidal *et al.*, Phys. Rev. A, vol. 62, 012304, 2000.
[8] M. A. Nielsen, Phys. Rev. Lett., vol. 83, pp. 436-439, 1999.