

## 量子系におけるベイズ予測の手法とその応用

東京大学・情報理工, 科学技術振興機構さきがけ (兼務)  
田中冬彦 (Fuyuhiko Tanaka)  
Graduate School of Information Science and Technology,  
The University of Tokyo and JST Sakigake

### Abstract

Tanaka and Komaki [17] では量子系でのベイズ予測が議論されていた. 測定方法を任意に固定した場合に, あるクラスの損失関数についてベイズの意味で最適な予測密度作用素が陽に与えられる. また, 測定方法も動かした場合にベイズの意味で最適な予測密度作用素を与える測定 (ベイズ測定と呼ぶ) の必要十分条件も紹介する.

## 1 Introduction

本研究では, 量子状態を表す密度作用素 (古典的な確率分布に対応) を推定することを考える. これまでの量子推定の研究では, 原理的に許されるあらゆる測定のもとでの最適なパラメータ推定方法を考えてきた [13, 14]. また, 標本数が十分大きい状況 (漸近的な状況) もしばしば考察の対象であった [10]. そのような場合には, 推定の理論的な精度限界などが詳しく調べられてきた. しかしながら, 現実的な状況では, むしろ, 測定装置は与えられていて, 未知のパラメータというよりも量子状態を表す密度作用素自身を知りたい場合も多い. これは物理の分野では状態再構成 (state reconstruction) などと呼ばれ, 非常に重要な話題である. 特に, 少数標本の場合には統計誤差が含まれるために, 再構成した密度作用素が正值作用素にはならないという問題が 1990 年代後半から認識されていた. Bužek *et al.* [6] らは, このような場合にはベイズ法が有効であることを主張していた. 彼らは, 古典的な事後分布の量子版を考えて, ベイズ予測密度作用素を用いるとうまくいくことを数値実験などで確認していた. しかしながら, なぜうまくいくのかということについて, 統計的な見地からの議論は全くなかった. Tanaka and Komaki [17] では, このような未知の密度作用素の推定を量子予測という枠組みで考察することを提唱し, ベイズ予測密度作用素があらゆる推定の中で最適になることを証明した. これは Aitchison [1] による古典ベイズ予測の基本的な結果の拡張である.

ただし, 上の証明では, 最適性をはかる基準として量子相対エントロピーのみを用いていた. 最適な予測方法というのは損失関数の取り方で一般には違ってくる. 本発表では, 量子  $\alpha$  ダイバージェンスという, 量子相対エントロピーや量子ヘルンジャー距離などを含む広いクラスの損失関数について, 最適な予測方法を提案する. 具体的には, 古典系と同様にして  $\alpha$  予測密度作用素を定義し, それが最適であることを示す. これらの結果は, Tanaka and Komaki [17] の結果を  $\alpha = -1$  の特殊ケー

スとして含んでいる。また、上記の議論は任意に測定方法を固定したもとの話であるが、原理的に許されるあらゆる測定のもとで平均損失を最小化する問題についても説明する。

なお、量子力学系における統計的推測の理論の必要性、歴史などは、林、松本による解説 [12] が詳しい。Hayashi [10] やレビュー論文 [5] も参考になる。

## 2 測定固定での量子ベイズ予測

Tanaka and Komaki [17] にしたがって量子系のベイズ予測を考えてみよう。基本的な問題設定は、Tanaka and Komaki [17] およびそこで引用されている文献を参照。 $\mathcal{H}$  は注目している物理系を記述するヒルベルト空間を表し、 $\mathcal{S}(\mathcal{H})$  は  $\mathcal{H}$  上の密度作用素の全体を表すものとする。ここでは、有限次元パラメータ  $\theta$  をもつ量子状態モデル  $\mathcal{M} := \{\rho_\theta \in \mathcal{S}(\mathcal{H}) | \theta \in \Theta \subset \mathbf{R}^k\}$  が与えられているものとする。さらに、情報抽出のための測定装置 (POVM)  $\{M_x\}$  も固定で考える。パラメータ  $\theta$  は  $\Theta$  上の絶対連続な事前分布  $\pi(\theta)d\theta$  に従うものとする。

ここでの目的は、測定で得られた結果  $x$  (確率分布  $p^M(x|\theta) = \text{Tr} M_x \rho_\theta$  にしたがって発生する) から、パラメータそのものを読み取るのではなく、量子状態  $\rho$  を推定するということである。(記法の簡便のため、確率分布は密度関数をもつとして表記する。離散の場合には単に積分  $\int dx$  を和記号  $\sum_x$  に置き換えるなどすればよい。)

古典予測が確率密度の推定と解釈できたように、量子予測は密度作用素の推定と考えるとよい。統計的決定理論の枠組みで最適な予測を議論するためには、真の密度作用素と推定した密度作用素との差をはかる指標が必要である。

### Definition 2.1

$\rho$  と  $\sigma$  を  $\mathcal{H}$  上の任意の密度作用素、 $\alpha \in \mathbf{R}$  とする。 $\rho$  から  $\sigma$  への量子  $\alpha$  ダイバージェンスは以下のように定義される。

$$D^{(\alpha)}(\rho||\sigma) := \frac{4}{1-\alpha^2} \left( 1 - \text{Tr} \sigma^{\frac{1+\alpha}{2}} \rho^{\frac{1-\alpha}{2}} \right), \quad \text{if } \alpha \neq 1$$

および

$$D^{(\alpha=-1)}(\rho||\sigma) := \text{Tr} \rho (\log \rho - \log \sigma) =: D^{(\alpha=1)}(\rho||\sigma).$$

### Remark 2.2

量子  $\alpha$  ダイバージェンスは  $\alpha$  ダイバージェンス ([2]) と同様に、対称ではないものの距離として望ましい性質を満たす。すなわち、 $D^{(\alpha)}(\rho||\sigma) \geq 0$  および  $D^{(\alpha)}(\rho||\sigma) = 0 \Leftrightarrow \rho = \sigma$ 。

### Remark 2.3

古典系と違いパラメータ  $\alpha$  は  $|\alpha| \leq 3$  でないと、結合凸性を満たさないことが知ら

れている. (See, Hasegawa [9] )

#### Remark 2.4

量子物理では, 以下のフィデリティ (fidelity) という量がよく使われる.

$$F(\rho, \sigma) := \text{Tr}|\sqrt{\rho}\sqrt{\sigma}| = \text{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}.$$

$0 \leq F \leq 1$  であり,  $F = 1 \leftrightarrow \rho = \sigma$  であるため,  $F$  が 1 に近いほど真の状態をよく推定していることになる. もし, 二つの密度作用素が純粋状態であれば,  $F(\rho, \sigma) = |\langle\varphi|\psi\rangle|$ ,  $\rho = |\varphi\rangle\langle\varphi|$  および  $\sigma = |\psi\rangle\langle\psi|$  と書ける. ブラケット記法やフィデリティの性質については例えば, Nielsen and Chuang [16], Hayashi [11].

さて, 通常の古典ベイズ統計と同じように事後分布を  $\pi(\theta|x) := \frac{\pi(\theta)p(x|\theta)}{\int d\theta\pi(\theta)p(x|\theta)}$  であらわそう. このとき,  $\pi(\theta|x)$  を利用して密度作用素の推定を考える.

#### Definition 2.5

$\alpha \in \mathbf{R}$  は任意に固定する.  $\alpha$  予測密度作用素は次のように定義される.

$$\tilde{\rho}_\pi^{(\alpha)}(x) := \frac{1}{B_\alpha(x)}\rho_\pi^{(\alpha)}(x), \quad B_\alpha(x) := \text{Tr}\rho_\pi^{(\alpha)}(x),$$

ただし,

$$\rho_\pi^{(\alpha)}(x) := \begin{cases} \{\rho_\theta^{\frac{1-\alpha}{2}}\pi(\theta|x)d\theta\}^{\frac{2}{1-\alpha}}, & \alpha \neq 1 \\ \exp\{\int \log(\rho_\theta)\pi(\theta|x)d\theta\}, & \alpha = 1. \end{cases}$$

#### Definition 2.6

$M$  粒子の  $\alpha$  予測密度作用素は次のように定義される.

$$\tilde{\sigma}_\pi^{(\alpha)}(x) := \frac{1}{C_\alpha(x)}\sigma_\pi^{(\alpha)}(x), \quad C_\alpha(x) := \text{Tr}\sigma_\pi^{(\alpha)}(x),$$

ここで,  $\sigma_\theta := \rho_\theta^{\otimes M}$  は  $M$  粒子全体のテンソル積状態を表し,

$$\sigma_\pi^{(\alpha)}(x) := \begin{cases} \{\int \sigma_\theta^{\frac{1-\alpha}{2}}\pi(\theta|x)d\theta\}^{\frac{2}{1-\alpha}}, & \alpha \neq 1, \\ \exp\{\int \log(\sigma_\theta)\pi(\theta|x)d\theta\}, & \alpha = 1. \end{cases}$$

である.

Corcuera and Giummolè [7] は古典的なベイズ予測の問題で一般化ベイズ予測密度  $p_\pi^{(\alpha)}(y|x)$  が,  $\alpha$  ダイバージェンスを損失関数にとるとき, 最適な予測密度であることを示した. 量子系でも同様な結果が成立する.

#### Theorem 2.7

任意に  $\alpha \in \mathbf{R}$  を固定する.  $N + M$  粒子系  $\rho_\theta^{\otimes(N+M)}$  において, 任意に選んだ  $N$  粒子

系  $\rho_\theta^{\otimes N}$  に対して測定を行い残りの  $M$  粒子の状態  $\sigma_\theta = \rho_\theta^{\otimes M}$  を推定するとしよう。真のパラメータの値  $\theta$  は未知であり、事前分布  $\pi(\theta)$  にしたがうと仮定する。任意の予測密度作用素を  $\hat{\sigma}(x)$  とあらわす。ここで  $x$  は  $N$  粒子系に対する測定  $\{M_x\}$  によって得られた値である。予測密度作用素の性能は、真の状態  $\sigma_\theta$  からの量子  $\alpha$  ダイバージェンスの平均

$$E^\pi E^{M_x}[D^{(\alpha)}(\sigma_\theta||\hat{\sigma})] = \int d\theta \pi(\theta) \int dx p(x|\theta) D^{(\alpha)}(\sigma_\theta||\hat{\sigma}(x))$$

で測ることとする。このとき、一般化ベイズ予測密度作用素  $\tilde{\sigma}_\pi^{(\alpha)}(x)$  が最良の予測密度作用素を与える。

*Proof.*

まず  $\alpha \neq \pm 1$  の時を考える。予測密度作用素のリスクの差  $D^{(\alpha)}(\sigma_\theta||\hat{\sigma}) - D^{(\alpha)}(\sigma_\theta||\tilde{\sigma}_\pi^{(\alpha)})$  の平均を計算してみる。ここからは、 $\tilde{\sigma}_\pi^{(\alpha)}$  の  $\alpha$  は省略することとする。

$$\begin{aligned} & E^\pi E^{M_x}[D^{(\alpha)}(\sigma_\theta||\hat{\sigma}) - D^{(\alpha)}(\sigma_\theta||\tilde{\sigma}_\pi)] \\ &= \int d\theta \pi(\theta) \int dx p(x|\theta) \left\{ \frac{4}{1-\alpha^2} \text{Tr} \sigma_\theta^{\frac{1-\alpha}{2}} (\tilde{\sigma}_\pi^{\frac{1+\alpha}{2}} - \hat{\sigma}^{\frac{1+\alpha}{2}}) \right\} \\ &= \int dx p_x \int d\theta \frac{\pi(\theta)p(x|\theta)}{p_x} \left\{ \frac{4}{1-\alpha^2} \text{Tr} \sigma_\theta^{\frac{1-\alpha}{2}} (\tilde{\sigma}_\pi^{\frac{1+\alpha}{2}} - \hat{\sigma}^{\frac{1+\alpha}{2}}) \right\} \\ &= \int dx p_x \int d\theta \pi(\theta|x) \left\{ \frac{4}{1-\alpha^2} \text{Tr} \sigma_\theta^{\frac{1-\alpha}{2}} (\tilde{\sigma}_\pi^{\frac{1+\alpha}{2}} - \hat{\sigma}^{\frac{1+\alpha}{2}}) \right\} \\ &= \int dx p_x \frac{4}{1-\alpha^2} \text{Tr} \left\{ \left( \int d\theta \pi(\theta|x) \sigma_\theta^{\frac{1-\alpha}{2}} \right) (\tilde{\sigma}_\pi^{\frac{1+\alpha}{2}} - \hat{\sigma}^{\frac{1+\alpha}{2}}) \right\} \\ &= \int dx p_x \frac{4}{1-\alpha^2} \text{Tr} \left\{ C_\alpha^{\frac{1-\alpha}{2}} \tilde{\sigma}_\pi^{\frac{1-\alpha}{2}} (\tilde{\sigma}_\pi^{\frac{1+\alpha}{2}} - \hat{\sigma}^{\frac{1+\alpha}{2}}) \right\} \\ &= \int dx p_x C_\alpha^{\frac{1-\alpha}{2}} \frac{4}{1-\alpha^2} \left\{ 1 - \text{Tr} \tilde{\sigma}_\pi^{\frac{1-\alpha}{2}} \hat{\sigma}^{\frac{1+\alpha}{2}} \right\} \\ &= \int dx p_x C_\alpha^{\frac{1-\alpha}{2}} D^{(\alpha)}(\tilde{\sigma}_\pi||\hat{\sigma}) \geq 0, \end{aligned}$$

ここで  $p_x := \int d\theta' \pi(\theta') p(x|\theta')$  は  $x$  のマージナルである。最後の不等式は、量子  $\alpha$  ダイバージェンスの正值性  $D^{(\alpha)}(\sigma||\sigma') \geq 0$  と  $p_x \geq 0$  からしたがう。上の式で、 $\hat{\sigma}(x)$  は任意の予測密度作用素であったから、 $\tilde{\sigma}_\pi^{(\alpha)}(x)$  は任意のほかの  $\hat{\sigma}(x)$  よりもリスクが小さい。これより、 $\tilde{\sigma}_\pi^{(\alpha)}(x)$  が最良であることが示された。同様に  $\alpha = \pm 1$  のときも示せる。

*Q.E.D.*

### Remark 2.8

以上の証明は、Hilbert-Schmidt ノルム ( $d_{\text{HS}}(\rho, \sigma)^2 := \text{Tr}[(\rho - \sigma)^2]$ ) を用いてもでき

る。その場合には、ベイズ予測密度作用素  $\sigma_\pi$  ( $\alpha = -1$  に対応) が最良である。

### Remark 2.9

フィデリティ損失の場合は,  $\text{spin}\frac{1}{2}$  の  $N$  粒子系  $\rho^{\otimes N}$  での 1 粒子予測は, すでに Bagan *et al.* [4] によって解かれている。ただし, 一般次元の場合, および,  $M$  粒子予測の場合にはまだ解かれていないようである。

以上は任意に測定を固定した場合の量子ベイズ予測の結果である。様々な計算例は著者の学位論文 [18] に詳しい。

## 3 測定の自由度も考慮した場合の量子ベイズ予測

次に測定を自由に選ぶことを考える。数学的に許される測定方法全体を動かした時に, ベイズ予測の平均リスクはどこまで小さくできるだろうか。特に最適な予測を与える測定方法が自然な形をとるのか, といったことは興味深い。個々の問題については以下の結果 [19] を利用して, 最適な測定方法と予測密度作用素の構成方法を示すことができる。

まず, 測定 (と密度作用素の推定) は  $\mathcal{S}(\mathcal{H})$  上に値をもつ以下のような正作用素値測度 (POVM) で表現できる。

$$\begin{aligned} M(B) &\geq 0, \forall B \in \mathcal{B}(\mathcal{S}(\mathcal{H})) \\ \sum M(B_j) &= M(\cup_j B_j), \quad \{B_j\} \subset \mathcal{B}(\mathcal{S}(\mathcal{H})) \text{ は disjoint} \\ M(\mathcal{S}(\mathcal{H})) &= I, \end{aligned}$$

ここで  $\mathcal{B}(\mathcal{S}(\mathcal{H}))$  は Borel 集合族である。このような正作用素値測度を  $\{M(d\hat{\rho})\}$  と表すことにする。  $\pi(d\theta)$  は  $\Theta$  上の proper な事前分布とする。  $w(\rho_\theta, \hat{\rho})$  で真の量子状態が  $\rho_\theta$  の時に  $\hat{\rho}$  と推定した時の損失を表すものとする。以降は  $w(\rho, \sigma)$  は  $\mathcal{S}(\mathcal{H}) \times \mathcal{S}(\mathcal{H})$  上有界連続な実数値関数とする。

### Definition 3.1

以下で定義される作用素を deviation operator とよぶ。

$$W(\hat{\rho}) := \int_{\Theta} \pi(d\theta) w(\rho_\theta, \hat{\rho}) \rho_\theta.$$

さらに,  $\mathcal{S}(\mathcal{H})$  上の作用素値関数  $T(\rho) : \mathcal{S}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H})$  に適当な可測性を定義 (See, e.g., Holevo [15]) することで

$$\int_{\mathcal{S}(\mathcal{H})} T(\rho) M(d\rho)$$

も通常の Lebesgue 積分と同様に定義できることに注意しておく.

### Theorem 3.2

$\dim \mathcal{H} < \infty$  とする. 上の表記にしたがうとき, POVM  $\{M^\circ(d\hat{\rho})\}$  について以下の条件は同値.

- (i)  $\{M^\circ(d\hat{\rho})\}$  はベイズ測定 (i.e., 平均リスク  $\mathcal{R}_\pi$  の最小値を達成)
- (ii) 自己共役作用素  $\Upsilon \in \mathcal{L}_h(\mathcal{H})$  が存在して,

$$\begin{aligned} \Upsilon &\leq W(\hat{\rho}), \quad \forall \hat{\rho} \in \mathcal{S}(\mathcal{H}) \\ (W(\hat{\rho}) - \Upsilon)M^\circ(d\hat{\rho}) &= 0, \quad \forall \hat{\rho} \in \mathcal{S}(\mathcal{H}) \end{aligned}$$

- (iii) 自己共役作用素  $\Upsilon \in \mathcal{L}_h(\mathcal{H})$  が存在して,

$$\begin{aligned} \Upsilon &\leq W(\hat{\rho}), \quad \forall \hat{\rho} \in \mathcal{S}(\mathcal{H}) \\ \Upsilon &= \int_{\mathcal{S}(\mathcal{H})} W(\hat{\rho})M^\circ(d\hat{\rho}) = \int_{\mathcal{S}(\mathcal{H})} M^\circ(d\hat{\rho})W(\hat{\rho}) \end{aligned}$$

### Theorem 3.3

$\dim \mathcal{H} < \infty$  とする. 未知の量子状態のパラメータ族  $\mathcal{M} = \{\rho_\theta\}$  が与えられているとする. 未知パラメータ  $\theta$  の Proper な事前分布を  $\pi(d\theta)$  とする. また, (1 粒子) 予測密度作用素  $\hat{\rho}$  を  $\mathcal{S}(\mathcal{H})$  にとることにして, POVM を  $\{M(d\hat{\rho})\}$  とする. このとき, 平均リスク

$$\mathcal{R}_\pi(\mathbf{M}) = \int_{\Theta} \pi(d\theta) \mathcal{R}(\theta, \mathbf{M})$$

where

$$\mathcal{R}(\theta, \mathbf{M}) = \int_{\mathcal{S}(\mathcal{H})} w(\rho_\theta, \hat{\rho}) \text{Tr} M(d\hat{\rho}) \rho_\theta$$

を最小とするような測定  $\{M_B\}$  が存在する.

### Remark 3.4

$N$  個の粒子  $\rho^{\otimes N} \in \mathcal{S}(\mathcal{H}^{\otimes N})$  に対して測定を行い,  $M$  個の粒子の合成状態  $\hat{\sigma} \in \mathcal{S}(\mathcal{H}^{\otimes M})$  を推定する場合にも直ちに拡張できる.

### Remark 3.5

量子物理, 特に量子情報に関連した分野では, ここで紹介した 1 粒子予測と実質, 同等のことを違う観点から議論している. 例えば, 量子ガウス状態におけるクローニングや量子メモリ, 量子鍵配送などではしばしば数学的には同等な問題が扱われており, 古典統計同様に予測の手法の応用範囲の広さを物語っている. 「ベイズ」

や「予測」といった単語を使っていないため、応用に関する文献を調べ上げるのは難しいが例えば, Bae and Acín [3], Hammerer *et al.* [8] をあげておく. これらは, フィデリティ損失でのベイズ予測の問題を取り扱っている.

## REFERENCES

- [1] J. Aitchison: Goodness of prediction fit. *Biometrika*, **62** (1975), 547-554.
- [2] S. Amari and H. Nagaoka: *Methods of Information Geometry*. AMS, Oxford, 2000.
- [3] J. Bae and A. Acín: Asymptotic quantum cloning is state estimation. *Phys. Rev. Lett.*, **97** (2006), 030402.
- [4] E. Bagan, M. Baig, R. Muñoz-Tapia, and A. Rodriguez: Collective versus local measurements in a qubit mixed-state estimation. *Phys. Rev. A*, **69** (2004), 010304(R).
- [5] O. E. Barndorff-Nielsen, R. D. Gill, and P. E. Jupp: On Quantum Statistical Inference. (with discussion) *J. Roy. Statist. Soc. Ser. B*, **65** (2003), 775–816.
- [6] V. Bužek, R. Derka, G. Adam and P. L. Knight: Reconstruction of quantum states of spin systems: from quantum Bayesian inference to quantum tomography. *Ann. Phys. (N.Y.)*, **266** (1998), 454-496.
- [7] J. M. Corcuera and F. Giummolè : A Generalized Bayes Rule for Prediction. *Scand. J. Statist.*, **26**, Issue 2, (1999), 265–279.
- [8] K. Hammerer, M. Wolf, E. Polzik, and J. Cirac: Quantum benchmark for storage and transmission of coherent states. *Phys. Rev. Lett.*, **94** (2005), 150503.
- [9] H. Hasegawa:  $\alpha$ -divergence of the non-commutative information geometry. *Rep. Math. Phys.*, **33** (1993), 87–93.
- [10] M. Hayashi: *Asymptotic Theory of Quantum Statistical Inference*. World Scientific, Singapore, 2005.
- [11] M. Hayashi: *Quantum Information: An Introduction*. Springer-Verlag, Berlin, 2006.
- [12] 林正人, 松本啓史: 量子系における統計的推測の最近の発展. 応用数理, vol. 11, No.3 (2001), 27-48.

- [13] C. W. Helstrom: *Quantum Detection Theory*. Academic Press, New York, 1976.
- [14] A. S. Holevo: *Probabilistic and Statistical Aspects of Quantum Theory*. North-Holland, Amsterdam, 1982.
- [15] A. S. Holevo: *Proc. of Steklov institute of Math.* **124** (1976), *AMS trans.* Issue 3 (1978).
- [16] M. A. Nielsen and I. L. Chuang: *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, 2000.
- [17] F. Tanaka and F. Komaki: Bayesian predictive density operators for exchangeable quantum-statistical models. *Phys. Rev. A*, **71** (2005), 052323.
- [18] F. Tanaka: *Geometrical approach to classical and quantum Bayesian prediction*, Ph.D. thesis, the University of Tokyo, Tokyo, 2007.
- [19] F. Tanaka: In preparation.