Inverse Problem and Emergence in Large Deviation Strategy

Izumi Ojima[A] and Kazuya Okamura[B]

[A]Research Institute for Mathematical Sciences, Kyoto University
[B]Department of Mathematics, Faculty of Science, Kyoto University

November 4, 2010

Contents

Who Guarantees Universality?

The most important notion in physics

Who Guarantees Universality?

The most important notion in physics = Universality

In standard theoretical framework, however, induction processes from measured data generated by the system of interest are not seriously taken into account, and hence, we do not know to which extent our theoretical hypotheses are realistic. Thus, if this difficulty is not resolved, no one can clarify what is the basis for universality.

Who Guarantees Universality?

The most important notion in physics = Universality

In standard theoretical framework, however, induction processes from measured data generated by the system of interest are not seriously taken into account, and hence, we do not know to which extent our theoretical hypotheses are realistic. Thus, if this difficulty is not resolved, no one can clarify what is the basis for universality.

In this study, therefore, we focus on statistical methods for developing induction.

Micro-Macro Duality

Micro-Macro Duality [Oj06] is a bidirectional method between deduction and induction, and can resolve the following dilemma.

Duheme-Quine thesis as a No-Go theorem [Oj10]

It is impossible to determine uniquely a theory from phenomenological data so as to reproduce the latter, because of unavoidable finiteness in number of measurable quantities and of their limited accuracy.

Using our strategy, deduction and induction, or "Micro" and "Macro", should be connected with each other by the idea of matching condition.

Large Deviation Strategy

Large Deviation Strategy (LDS)
= Step-by-Step method of induction based on Large Deviation Principle
+ Micro-Macro duality formulated in the quadrality scheme [Oj10]
consisting of the following four basic ingredients:

1. Algebra ($\boldsymbol{Alg}$)
2. States ($\boldsymbol{States}$) and Representations ($\boldsymbol{Reps}$)
3. Spectrum ($\boldsymbol{Spec}$)
4. Dynamics ($\boldsymbol{Dyn}$)

LDS is based on the following four levels.

1st level : Abelian von Neumann algebras

Gel'fand rep., Strong law of large numbers(SLLN) and statistical inference on abelian v.N. alg.

2nd level : $\boldsymbol{States}$ and $\boldsymbol{Reps}$

Measure-theoretical analysis for noncommutative algebras

3rd level : $\boldsymbol{Spec}$ and $\boldsymbol{Alg}$

Emergence of space-time and composite system

4th level : $\boldsymbol{Dyn}$

From emergence to space-time patterns and time-series analysis

Several methods which play central roles in LDS

I. Large deviation principle [DS,E]
   From probablistic fluctuation and statistical inference

II. Tomita decomposition theorem and central decomposition
   How to formulate and use state-valued random variables

III. The dual $\widehat{G}$ of a group G and its crossed products
   From Macro to Micro

IV. Emergence : Condensation associated with spontaneous symmetry breaking(SSB) and phase separation
   From Micro to Macro

V. Operator-valued kernel method

1st level: Abelian von Neumann Algebras

Let $\mathcal{A}$ be an abelian v.N. alg. and ω be a normal state on $\mathcal{A}$. It holds that

$$\langle \Omega_\omega, \pi_\omega(A)\Omega_\omega \rangle = \omega(A) = \int \hat{A}(k) d\nu_\omega(k), \tag{1}$$

$$\pi_\omega(\mathcal{A}) \cong L^\infty(K, \nu_\omega), \quad \mathfrak{H}_\omega \cong L^2(K, \nu_\omega), \tag{2}$$

$$\mathcal{A}_* \cong L^1(K, \nu_\omega), \quad \Omega_\omega \leftrightarrow 1, \tag{3}$$

where K is a compact Hausdorff space and ν_ω is a Borel measure on K.

1st level: Abelian von Neumann Algebras

Let $\mathcal{A}$ be an abelian v.N. alg. and ω be a normal state on $\mathcal{A}$. It holds that

$$\langle \Omega_\omega, \pi_\omega(A)\Omega_\omega \rangle = \omega(A) = \int \hat{A}(k) d\nu_\omega(k), \tag{1}$$

$$\pi_\omega(\mathcal{A}) \cong L^\infty(K, \nu_\omega), \quad \mathfrak{H}_\omega \cong L^2(K, \nu_\omega), \tag{2}$$

$$\mathcal{A}_* \cong L^1(K, \nu_\omega), \quad \Omega_\omega \leftrightarrow 1, \tag{3}$$

where K is a compact Hausdorff space and ν_ω is a Borel measure on K.

Every self-adjoint element $\pi_\omega(A)$ of $\pi_\omega(\mathcal{A})$ is treated as measure-theoretical $\mathbb{R}$-valued random variable $\hat{A}$. Thus, we can discuss spectra of observables in the commutative case.

For any $\bar{k} = (k_1, k_2, \cdots) \in K^{\mathbb{N}}$ and $A = A^* \in \mathcal{A}$, we define $X_j(\bar{k}) = k_j$ and $\hat{A}_j(\bar{k}) := \hat{A}(X_j(\bar{k}))$.

Matching condition 1

$\{\hat{A}_j\}$ are independent identically distributed ("i.i.d.") random variables.

For any $\bar{k} = (k_1, k_2, \cdots) \in K^{\mathbb{N}}$ and $A = A^* \in \mathcal{A}$, we define $X_j(\bar{k}) = k_j$ and $\hat{A}_j(\bar{k}) := \hat{A}(X_j(\bar{k}))$.

Matching condition 1

$\{\hat{A}_j\}$ are independent identically distributed (“i.i.d.”) random variables.

Cramér's theorem

Let $M_n(\bar{k}) = \frac{1}{n}(\hat{A}_1(\bar{k}) + \cdots + \hat{A}_n(\bar{k}))$ and $Q_n^{(1)}(\Gamma) = P_{\nu_\omega}(M_n \in \Gamma)$.Then, $Q_n^{(1)}$ satisfies LDP with the rate function $I_\omega(a)$ $= \sup_{t\in\mathbb{R}}\{at - c_\omega(t)\}$ $(c_\omega(t) = \log \int_{\mathbb{R}} e^{tx}\nu_\omega(\hat{A} \in dx))$:

$$
\begin{aligned}
-\inf_{a\in\Gamma^o} I_\omega(a) &\leq \liminf_{n\to\infty} \frac{1}{n} \log Q_n^{(1)}(\Gamma) \\
&\leq \limsup_{n\to\infty} \frac{1}{n} \log Q_n^{(1)}(\Gamma) \leq -\inf_{a\in\overline{\Gamma}} I_\omega(a) \qquad (4)
\end{aligned}
$$

Other analysis are the same methods used in 2nd level and are omitted here.

2nd level: States and Representations

LDS 2nd level goes through the following procedures.

1. Tomita decomposition theorem and central decomposition
2. Sanov's theorem and quantum relative entropy as rate function
3. Bayesian escort predictive state
4. Singular statistics

2nd level: States and Representations

LDS 2nd level goes through the following procedures.

1. Tomita decomposition theorem and central decomposition
2. Sanov's theorem and quantum relative entropy as rate function
3. Bayesian escort predictive state
4. Singular statistics

The notion of sectors is cruicial for LDS 2nd and 3rd levels.

Definition 1. (Sector)

A sector of an algebra $\mathfrak{A}$ is defined by a quasi-equivalence class of factor states of $\mathfrak{A}$.

Two representations π_1 and π_2 are quasi-equivalent $\pi_1 \approx \pi_2$ if π_1 and π_2 are unitary equivalent up to multiplicity. Each sector corresponds to a pure phase parametrized by a spectrum $\eta \in \mathbf{Spec}\mathfrak{Z}_\omega(\mathfrak{A})$ of the order parameters constituting the center $\mathfrak{Z}_\omega(\mathfrak{A}) = \pi_\omega(\mathfrak{A})'' \cap \pi_\omega(\mathfrak{A})'$ of $\pi_\omega(\mathfrak{A})''$.

Tomita decomposition theorem (see [BR])

Let $\mathfrak{A}$ be a C*-algebra and ω be a state on $\mathfrak{A}$. There is a one-to-one correspondence between the following three sets.

(i) the orthogonal measure$^{(*)}$ μ $(\in \mathcal{O}_\omega(E_{\mathfrak{A}}))$ on $E_{\mathfrak{A}}$ with barycenter ω;

(ii) the abelian v.N. subalgebra $\mathfrak{B} \subseteq \pi_\omega(\mathfrak{A})'$;

(iii) the projection operator P on $\mathfrak{H}_\omega$ such that

$$P\Omega_\omega = \Omega_\omega, \quad P\pi_\omega(\mathfrak{A})P \subset \{P\pi_\omega(\mathfrak{A})P\}'.$$

If $\mu, \mathfrak{B}, P$ are in correspondence one has the following relation.

$\mathfrak{B}$ is $*$-isomorphic to the map $L^\infty(\mu) \ni f \mapsto \kappa_\mu(f) \in \pi_\omega(\mathfrak{A})'$ defined by

$$\langle \Omega_\omega, \kappa_\mu(f)\pi_\omega(A)\Omega_\omega\rangle = \int d\mu(\omega')f(\omega')\widehat{A}(\omega')$$

and for $A, B \in \mathfrak{A}$

$$\kappa_\mu(\hat{A})\pi_\omega(B)\Omega_\omega = \pi_\omega(B)P\pi_\omega(A)\Omega_\omega.$$

$(*)$ Orthogonality in the sense that

$$\int_\Delta \chi d\mu(\chi) \perp \int_{E_{\mathfrak{A}} \setminus \Delta} \chi d\mu(\chi)$$

for every $\Delta \in \mathcal{B}(E_{\mathfrak{A}})$.

For the map $\mathfrak{A} \ni A \longmapsto \widehat{A} \in L^\infty(\mu)$ defined by $\widehat{A} := (E_{\mathfrak{A}} \ni \omega \longmapsto \omega(A))$, its image $\widehat{\mathfrak{A}} := \{\widehat{A} | A \in \mathfrak{A}\}$ constitutes a C*-bialgebra of measure-theoretical random variables w.r.t. a linear structure $(\alpha\widehat{A} + \beta\widehat{B})(\omega) := (\alpha\widehat{A + \beta}B)(\omega)$ $(\alpha, \beta \in \mathbb{C})$, a commutative and a non-commutative convolution products defined by $(\widehat{A} \cdot \widehat{B})(\omega)$ $:= \widehat{A}(\omega)\widehat{B}(\omega)$ and $(\widehat{A} * \widehat{B})(\omega) := \widehat{AB}(\omega)$ with the norm $\|\cdot\|$ given by $\|\widehat{A}\| = \sup_{\omega \in E_{\mathfrak{A}}, \|\omega\|=1} |\widehat{A}(\omega)|$.

For the map $\mathfrak{A} \ni A \longmapsto \widehat{A} \in L^\infty(\mu)$ defined by $\widehat{A} := (E_\mathfrak{A} \ni \omega \longmapsto \omega(A))$, its image $\widehat{\mathfrak{A}} := \{\widehat{A} | A \in \mathfrak{A}\}$ constitutes a C*-bialgebra of measure-theoretical random variables w.r.t. a linear structure $(\alpha\widehat{A} + \beta\widehat{B})(\omega) := (\widehat{\alpha A + \beta B})(\omega)$ $(\alpha, \beta \in \mathbb{C})$, a commutative and a non-commutative convolution products defined by $(\widehat{A} \cdot \widehat{B})(\omega)$ $:= \widehat{A}(\omega)\widehat{B}(\omega)$ and $(\widehat{A} * \widehat{B})(\omega) := \widehat{AB}(\omega)$ with the norm $\|\cdot\|$ given by $\|\widehat{A}\| = \sup_{\omega \in E_\mathfrak{A}, \|\omega\|=1} |\widehat{A}(\omega)|$.

Definition 2. (Central and subcentral measures)

The measure μ $(= \mu_\mathfrak{B})$ $\in \mathcal{O}_\omega(E_\mathfrak{A})$ is called a subcentral measure of ω, if the algebra $\mathfrak{B}$ corresponding to μ is a subalgebra of the center $\mathfrak{Z}_\omega(\mathfrak{A})$. In particular, the subcentral measure $\mu_{\mathfrak{Z}_\omega(\mathfrak{A})} =: \mu_\omega$ is called a central measure of ω.

For the map $\mathfrak{A} \ni A \longmapsto \widehat{A} \in L^{\infty}(\mu)$ defined by $\widehat{A} := (E_{\mathfrak{A}} \ni \omega \longmapsto \omega(A))$, its image $\widehat{\mathfrak{A}} := \{\widehat{A} | A \in \mathfrak{A}\}$ constitutes a C*-bialgebra of measure-theoretical random variables w.r.t. a linear structure $(\alpha\widehat{A} + \beta\widehat{B})(\omega) := (\widehat{\alpha A + \beta B})(\omega)$ $(\alpha, \beta \in \mathbb{C})$, a commutative and a non-commutative convolution products defined by $(\widehat{A} \cdot \widehat{B})(\omega)$ $:= \widehat{A}(\omega)\widehat{B}(\omega)$ and $(\widehat{A} * \widehat{B})(\omega) := \widehat{AB}(\omega)$ with the norm $\| \cdot \|$ given by $\|\widehat{A}\| = \sup_{\omega \in E_{\mathfrak{A}}, \|\omega\|=1} |\widehat{A}(\omega)|$.

Definition 2. (Central and subcentral measures)

The measure μ $(= \mu_{\mathfrak{B}}) \in \mathcal{O}_{\omega}(E_{\mathfrak{A}})$ is called a subcentral measure of ω, if the algebra $\mathfrak{B}$ corresponding to μ is a subalgebra of the center $\mathfrak{Z}_{\omega}(\mathfrak{A})$. In particular, the subcentral measure $\mu_{\mathfrak{Z}_{\omega}(\mathfrak{A})} =: \mu_{\omega}$ is called a central measure of ω.

The set $\{\kappa_{\mu_{\omega}}(\chi_{\Delta}) | \Delta \in \mathcal{B}(\text{supp } \mu_{\omega})\}$ forms a projection-valued measure $E_{\omega} := (\mathcal{B}(\text{supp } \mu_{\omega}) \ni \Delta \mapsto E_{\omega}(\chi_{\Delta}) := \kappa_{\mu_{\omega}}(\chi_{\Delta})$ $\in \mathfrak{Z}_{\omega}(\mathfrak{A}))$ satisfiying

$$\langle \Omega_{\omega}, E_{\omega}(\Delta)\Omega_{\omega} \rangle = \langle \Omega_{\omega}, \kappa_{\mu_{\omega}}(\chi_{\Delta})\Omega_{\omega} \rangle = \mu_{\omega}(\Delta).$$

Suppose that $\mathfrak{A}$ is separable.

For $\tilde{\omega} = (\omega_1, \omega_2, \cdots) \in (\mathrm{supp}\ \mu_\omega)^{\mathbb{N}}$, $A \in \mathcal{B}(\mathrm{supp}\ \mu_\omega)$ and $\Gamma \in \mathcal{B}(E_{\mathfrak{A}})$, we define

$$Y_j(\tilde{\omega}) = \omega_j,$$
$$L_n(\tilde{\omega}, A) = \frac{1}{n}\sum_{j=1}^{n} \delta_{Y_j(\tilde{\omega})}(A),$$
$$Q_n^{(2)}(\Gamma) = P_{\mu_\omega}(L_n \in \Gamma).$$

Suppose that $\mathfrak{A}$ is separable.

For $\tilde{\omega} = (\omega_1, \omega_2, \cdots) \in (\mathbf{supp}\ \mu_\omega)^{\mathbb{N}}$, $A \in \mathcal{B}(\mathbf{supp}\ \mu_\omega)$ and $\Gamma \in \mathcal{B}(E_{\mathfrak{A}})$, we define

$$Y_j(\tilde{\omega}) = \omega_j,$$
$$L_n(\tilde{\omega}, A) = \frac{1}{n} \sum_{j=1}^{n} \delta_{Y_j(\tilde{\omega})}(A),$$
$$Q_n^{(2)}(\Gamma) = P_{\mu_\omega}(L_n \in \Gamma).$$

Matching condition 2.

$\{Y_j\}$ are independent identically distributed ("i.i.d.") random variables.

Suppose that $\mathfrak{A}$ is separable.
For $\tilde{\omega} = (\omega_1, \omega_2, \cdots) \in (\mathbf{supp}\ \mu_\omega)^{\mathbb{N}}$, $A \in \mathcal{B}(\mathbf{supp}\ \mu_\omega)$ and $\Gamma \in \mathcal{B}(E_{\mathfrak{A}})$, we define

$$Y_j(\tilde{\omega}) = \omega_j,$$
$$L_n(\tilde{\omega}, A) = \frac{1}{n} \sum_{j=1}^{n} \delta_{Y_j(\tilde{\omega})}(A),$$
$$Q_n^{(2)}(\Gamma) = P_{\mu_\omega}(L_n \in \Gamma).$$

Matching condition 2.

$\{Y_j\}$ are independent identically distributed ("i.i.d.") random variables.

The next theorem [HOT83] is the key to proving LDP.

Theorem 2.

Let μ, ν be regular Borel probability measures on $E_{\mathfrak{A}}$ with barycenters $\psi, \omega \in E_{\mathfrak{A}}$. If there is a subcentral measure m on $E_{\mathfrak{A}}$ such that $\mu, \nu \ll m$, then $S(\psi \| \omega) = D(\mu \| \nu)$.

Sanov's theorem

If there exists a subcentral measure m on $E_{\mathfrak{A}}$ such that $\mu_\omega \ll m$, then $Q_n^{(2)}$ satisfies LDP with the rate function $S(\cdot\|\omega)$:

$$\begin{aligned} -\inf_{\substack{\psi\in\Gamma^o,\\ \mu_\psi\ll\mu_\omega}} S(\psi\|\omega) &\le \liminf_{n\to\infty}\frac{1}{n}\log Q_n^{(2)}(\Gamma) \\ &\le \limsup_{n\to\infty}\frac{1}{n}\log Q_n^{(2)}(\Gamma) \le -\inf_{\substack{\psi\in\overline{\Gamma},\\ \mu_\psi\ll\mu_\omega}} S(\psi\|\omega) \end{aligned} \tag{5}$$

Definition 3. (model)

A family of states $\{\omega_\lambda | \lambda \in \Lambda \subset \mathbb{R}\text{: cpt}\}$ is called a (statistical) model if it satisfies the following two conditions.

(i) There is a subcentral measure m on $E_{\mathfrak{A}}$ such that $\mu_{\omega_\lambda} \ll m$ for every $\lambda \in \Lambda$.

(ii) The set $\overline{\{\rho \in E_{\mathfrak{A}} | \frac{d\mu_{\omega_\lambda}}{dm}(\rho) > 0\}}$ does not depend on $\lambda \in \Lambda$.

Definition 3. (model)

A family of states $\{\omega_\lambda | \lambda \in \Lambda \subset \mathbb{R}: \mathbf{cpt}\}$ is called a (statistical) model if it satisfies the following two conditions.

(i) There is a subcentral measure m on $E_{\mathfrak{A}}$ such that $\mu_{\omega_\lambda} \ll m$ for every $\lambda \in \Lambda$.

(ii) The set $\overline{\{\rho \in E_{\mathfrak{A}} | \frac{d\mu_{\omega_\lambda}}{dm}(\rho) > 0\}}$ does not depend on $\lambda \in \Lambda$.

Definition 4. (Bayesian escort predictive state)

Let $\{\omega_\lambda\}_{\lambda \in \Lambda}$ be a model, $\pi(\lambda)$ be a probability distribution on Λ, $p_\lambda(x)$ be a probability distribution dependent on $\{\omega_\lambda\}_{\lambda \in \Lambda}$, $x^n = \{x_1, \cdots, x_n\}$ and $\beta > 0$. The state

$$\omega^{x^n}_{\pi,\beta} = \frac{\int \omega_\lambda \prod_{j=1}^n p_\lambda(x_j)^\beta \pi(\lambda) d\lambda}{\int \prod_{j=1}^n p_\lambda(x_j)^\beta \pi(\lambda) d\lambda} \tag{6}$$

is called a Bayesian escort predictive state. When p_λ is equal to $\frac{d\mu_{\omega_\lambda}}{dm}$, we write $\omega^n_{\pi,\beta} = \omega^{\rho^n}_{\pi,\beta}$.

Theorem 3.

The risk function

$$\boldsymbol{T^n(\psi^{x^n}\|\omega_\lambda) = \frac{1}{A}\iint S(\psi^{x^n}\|\omega_\lambda)\prod_{j=1}^{n} p_\lambda(x_j)^\beta d\nu(x_j)\pi(\lambda)d\lambda,} \tag{7}$$

$$\boldsymbol{A = \iint \prod_{j=1}^{n} p_\lambda(x_j)^\beta d\nu(x_j)\pi(\lambda)d\lambda}$$

of ψ^{x^n},which depends on data x^n, is minimized at the Bayesian escort predictive state $\omega_{\pi,\beta}^{x^n}$.

This result is a generalization of [Ai75] and [TK05], and is the reason that the Bayesian escort predictive state is a good estimator for “true” one.

Now we discuss singular statistics. The results here are proved originally in [W,W10]. Let $\{\omega_\lambda\}_{\lambda\in\Lambda}$ be a model and $\phi \in E_{\mathfrak{A}}$ such that there is a subcentral measure m satisfying $\mu_{\omega_\lambda}, \mu_\phi \ll m$ and $\overline{\{\rho \in E_{\mathfrak{A}} | \, p_\lambda(\rho) := \frac{d\mu_{\omega_\lambda}}{dm}(\rho) > 0\}} = \overline{\{\rho \in E_{\mathfrak{A}} | \, q(\rho) := \frac{d\mu_\phi}{dm}(\rho) > 0\}}$ for every $\lambda \in \Lambda$.

Now we discuss singular statistics. The results here are proved originally in [W,W10]. Let $\{\omega_\lambda\}_{\lambda\in\Lambda}$ be a model and $\phi \in E_{\mathfrak{A}}$ such that there is a subcentral measure m satisfying $\mu_{\omega_\lambda}, \mu_\phi \ll m$ and $\{\rho \in E_{\mathfrak{A}} |\ p_\lambda(\rho) := \frac{d\mu_{\omega_\lambda}}{dm}(\rho) > 0\} = \{\rho \in E_{\mathfrak{A}} |\ q(\rho) := \frac{d\mu_\phi}{dm}(\rho) > 0\}$ for every $\lambda \in \Lambda$.

Definition 5. (Partition function and Likelihood)

$$Z_n = \int \prod_{j=1}^{n} p(\rho_j|\lambda)^\beta \pi(\lambda) d\lambda, \qquad Z_n^0 = \frac{Z_n}{\prod_{j=1}^{n} q(\rho_j)^\beta}, \tag{8}$$

$$F_n = -\frac{1}{\beta}\log Z_n, \qquad F_n^0 = -\frac{1}{\beta}\log Z_n^0. \tag{9}$$

Z_n and $\frac{1}{n}F_n$ is called a partition function and an empirical logarithmic Bayesian escort likelihood, respectively.

Theorem 4.

Let $D(\lambda) = D(p_\lambda \| q)$ $(D(0) = 0)$, $f(\rho, \lambda) = \log \frac{q(\rho)}{p_\lambda(\rho)}$ and $D_n(\lambda) = \frac{1}{n} \sum_{j=1}^{n} f(\rho_j, \omega)$. By resolution of singularities, it holds that

$$
\begin{aligned}
D(g(u)) &= u^{2k} = u_1^{2k_1} \dots u_d^{2k_d}, && (10)\\
f(\rho, g(u)) &= a(\rho, u) u^k, && (11)\\
D_n(g(u)) &= u^{2k} - \frac{1}{\sqrt{n}} u^k \xi_n(u), && (12)
\end{aligned}
$$

where $u = (u_1, \cdots, u_d)$ is a coordinate system of an analytic manifold U, and g is an analytic map from U to Λ, $k_1, \cdots, k_d$ are non-negative integers, $a(\rho, u)$ is an analytic function on U for each $\rho \in \mathrm{supp}\ \mu_{\omega_\lambda}$, and $\{\xi_n\}$ is an empirical process such that

$$
\xi_n(u) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \{a(\rho_j, u) - u^k\}, \qquad (13)
$$

which converges to a gaussian process $\xi(u)$ weakly.

The zeta function $\zeta(z) = \int_{C:\mathrm{cpt}} D(\lambda)^z \pi(\lambda) d\lambda$ can be analytically continued to the unique meromorphic function on the entire complex plane. All poles of $\zeta(z)$ are real, negative, rational numbers.

$$(-\lambda) := \text{maximum poles of } \zeta(z) \ (\lambda > 0),$$
$$m := \text{multiplicity of } (-\lambda).$$

Theorem 5.

(1)

$$F_n^0 - \frac{\lambda}{\beta}\log n + \frac{m-1}{\beta}\log\log n$$
$$\longrightarrow -\frac{1}{\beta}\log\left(\sum_{\alpha^*}\gamma_b \int_0^\infty dt \int t^{\lambda-1} e^{-\beta t + \beta\sqrt{t}\xi_0(y)} \varphi_0^*(y) dy\right). \tag{14}$$

(2)

$$\frac{1}{n}F_n - \frac{\lambda}{\beta}\frac{\log n}{n} + \frac{m-1}{\beta}\frac{\log\log n}{n} \longrightarrow S, \tag{15}$$

where $S = -\int dm(\rho) q(\rho) \log q(\rho)$ is the Shannon entropy.

We define, for $\rho^n = \{\rho_1, \cdots, \rho_n\}$,

$$\langle f(\lambda)\rangle_{\pi,\beta}^{\rho^n} = \frac{\int f(\lambda)\prod_{j=1}^n p_\lambda(\rho_j)^\beta \pi(\lambda)d\lambda}{\int \prod_{j=1}^n p_\lambda(\rho_j)^\beta \pi(\lambda)d\lambda}. \tag{16}$$

Definition 6. (Errors, Losses and variance)

(1) Bayes generalization error (loss),

$$B_g = E_\rho\left[\log \frac{q(\rho)}{\langle p_\lambda(\rho)\rangle_{\pi,\beta}^{\rho^n}}\right], BL_g = E_\rho\left[-\log\langle p_\lambda(\rho)\rangle_{\pi,\beta}^{\rho^n}\right],$$

respectively.

(2) Bayes training error (loss),

$$B_t = \frac{1}{n}\sum_{j=1}^n \left[\log \frac{q(\rho_j)}{\langle p_\lambda(\rho_j)\rangle_{\pi,\beta}^{\rho^n}}\right], BL_t = \frac{1}{n}\sum_{j=1}^n \left[-\log\langle p_\lambda(\rho_j)\rangle_{\pi,\beta}^{\rho^n}\right], \tag{17}$$

respectively.

(3) functional variance,

$$V = \sum_{j=1}^n \left\{\langle p_\lambda(\rho_j)^2\rangle_{\pi,\beta}^{\rho^n} - (\langle p_\lambda(\rho_j)\rangle_{\pi,\beta}^{\rho^n})^2\right\}.$$

Let $\Lambda_\epsilon = \{\lambda \in \Lambda | D(p_\lambda \| p_0) \leq \epsilon\}$. If there exists $A > 0$ and $\epsilon > 0$ such that $\lambda \in \Lambda_\epsilon \;\Rightarrow\; \int dm(\rho) q(\rho) \log \frac{p_0(\rho)}{p_\lambda(\rho)} \geq A \cdot D(p_\lambda \| p_0)$, then the pair (p_λ, q) is said to be coherent.

Theorem 6.

If the pair (p_λ, q) satisfies the coherence condition, then it holds that

$$E_{\rho^n}[BL_g] = E_{\rho^n}[\mathrm{WAIC}] + o\left(\frac{1}{n}\right), \quad (18)$$

$$\mathrm{WAIC} = BL_t + \frac{\beta}{n} V. \quad (19)$$

Let $\Lambda_\epsilon = \{\lambda \in \Lambda | D(p_\lambda \| p_0) \leq \epsilon\}$. If there exists $A > 0$ and $\epsilon > 0$ such that $\lambda \in \Lambda_\epsilon \Rightarrow \int dm(\rho) q(\rho) \log \frac{p_0(\rho)}{p_\lambda(\rho)} \geq A \cdot D(p_\lambda \| p_0)$, then the pair (p_λ, q) is said to be coherent.

Theorem 6.

If the pair (p_λ, q) satisfies the coherence condition, then it holds that

$$E_{\rho^n}[BL_g] = E_{\rho^n}[\mathrm{WAIC}] + o\left(\frac{1}{n}\right), \quad (18)$$

$$\mathrm{WAIC} = BL_t + \frac{\beta}{n} V. \quad (19)$$

Since WAIC for $p_\lambda = \frac{d\mu_{\omega_\lambda}}{dm}$ is a quantum version of the information criteria (IC), we can successfully interpret this result as establishing IC for quantum states. This also justifies our use of the central measure μ_ω of $\omega \in E_{\mathfrak{A}}$. On the other hand, IC in the 1st level are the same as those in classical case. In practical situations to use the methods discussed in this section, it will be safe for them to be applied to only the case:

$$\omega_\lambda = \int \rho \, d\mu_{\omega_\lambda}(\rho) = \int_B \rho_\xi \, d\tilde{\mu}_\lambda(\xi)$$

where $\{\rho_\xi | \xi \in \Xi : \textbf{an order parameter}\} \subset F_{\mathfrak{A}}$, and B is compact.

Conclusion and Perspective

We have established LDS 1st and 2nd level.

In contrast, LDS 3rd and 4th levels remain to be analyzed. In details,

- How to identify the Spec of order parameters with a homogeneous space G/H?

- How can we reconstruct the total algebra $\mathcal{F}$ acted on by a group G from the G-fixed observable algebra $\mathfrak{A}$?

- Several kernels $\{K_\tau(\cdot,\cdot)\}_{\tau\in T}$ are strongly related to the action α of the locally compact group G.

$$Aut(G) \ni \alpha \longleftrightarrow \{K_\tau(\cdot,\cdot)\}_{\tau\in T}$$

to be continued...

Reference

[Oj06] I. Ojima,"Micro-Macro Duality in Quantum Physics", pp.143-161 in Proc. Intern. Conf. on Stochastic Analysis, Classical and Quantum (World Scientific, 2005), arXiv:math-ph/0502038.

[Oj10] I. Ojima, J. Phys.: Conf. Ser. **201**, 012017 (2010).

[DS] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications* 2nd eds. (Springer, 1997).

[E] R. S. Ellis, *Entropy, Large Deviations, and Statistical Mechanics*, (Springer, 1985).

[Oj98] I. Ojima, Order Parameters in QFT and Large Deviation, RIMS Kokyuroku **1066** 121-132 (1998), (in Japanese), http://repository.kulib.kyoto-u.ac.jp/dspace/bitstream/2433/62481/1/1066-10.pdf.

[BR] O. Bratteli and D. W. Robinson, *Operator algebras and Quantum Statistical Mechanics* vol.1 (Springer, 1979).

[HOT83] F. Hiai, M. Ohya and M. Tsukada, Pacific J. Math. **107**, 117-140 (1983).

Reference : continued

[Ai75] J. Aitchison, Biometrika **62**, 547 (1975).

[KT05] F. Komaki and F. Tanaka, Phys. Rev. A **71**, 052323 (2005).

[W] S. Watanabe, *Algebraic geometry and statistical learning theory*, (Cambridge University Press, 2009).

[W10] S. Watanabe, J. Phys.: Conf. Ser. **233**, 012014 (2010).

[NT] Y. Nakagami and M. Takesaki, *Duality for Crossed Products of von Neumann Algebras*, Lec. Notes in Math. **731**, (Springer, 1979).

[IT78] S. Imai and H. Takai, J. Math. Soc. Japan **30**, 495-504 (1978).

[Oj03] I. Ojima, Open Sys. Info. Dyn. **10**, 235-279 (2003).

[Oj10] I. Ojima, "Dilation and Emergence in Physical Sciences", Invited talk at Int. Conf., "Advances in Quantum Theory" at Linnaeus Univ., June 2010.

[OjOz93] I. Ojima and M. Ozawa, Open Sys. Info. Dyn. **2**, 107 (1993).

[OjOk] I. Ojima and K. Okamura, in preparation.