みんなで学ぶ数理物理 (大偏差原理)

SG2023-10

1 General Theory of LDP

Let X be a Polish space.

Definition (rate function)
A function $I : X \to [0,\infty]$ is called a rate function if
(i) $I \not\equiv \infty$.
(ii) I has compact level sets, i.e., $f^{-1}([-\infty, c]) = \{x \in X : f(x) \leq c\}$ is compact for all $c \in \mathbb{R}$.

For a rate function I and subset $S \subset X$, we define

$$I(S) := \inf_{x \in S} I(x).$$

Definition (Large Deviation Principle)
A sequence of probability measures (P_n) on X is said to satisfy the *large deviation principle* (LDP) with rate n and with rate function I if
(i) I is a rate fimction.
(ii) $\limsup_{n\to\infty} \frac{1}{n} \log P_n(C) \leq -I(C) \quad \forall C \subset X$ closed.
(iii) $\liminf_{n\to\infty} \frac{1}{n} \log P_n(O) \geq -I(O) \quad \forall O \subset X$ open.

Let (P_n) satisfy the LDP on X with rate n and with rate function I. Let $F : X \to \mathbb{R}$ be a continuous function that is bounded from above.

Theorem (Varadhan's Lemma)

$$\lim_{n\to\infty} \frac{1}{n} \log \int_X e^{nF(x)} P_n(dx) = \sup_{x\in X} [F(x) - I(x)].$$

Theorem (Tilted LDP)
We define $J_n(S) := \int_S e^{nF(x)} P_n(dx)$ ($S \subset X$ Borel). Then the sequence (P_n^F) defined by

$$P_n^F(S) := \frac{J_n(S)}{J_n(X)}, \quad S \subset X \text{ Borel},$$

satisfies the LDP with rate n and with rate function

$$I^F(x) := \sup_{y\in X} [F(y) - I(y)] - [F(x) - I(x)].$$

Theorem (Contraction Principle)
Let Y be a Polish space and $T : X \to Y$ be a continuous map. Then the sequence of image probability measures $(Q_n) := (P_n \circ T^{-1})$ satisfies the LDP on Y with rate n and with rate function J given by

$$J(y) := \inf_{x\in X:\, T(x)=y} I(x).$$

2 LDP for I.I.D. sequences

We assume that X_i take values in a finite set:

(i) $X_i \in \Gamma = \{1, 2, \ldots, r\} \subset \mathbb{N}$.

(ii) $X_1, X_2, \cdots$ are i.i.d. with marginal law $\rho = (\rho_s)_{s\in\Gamma}$, i.e., $\mathbb{P}(X_i = s) = \rho_s$, $\forall s \in \Gamma$.

(iii) $\rho_s > 0$, $\forall s \in \Gamma$.

We introduce the empirical measure

$$L_n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i},$$

with δ_x denoting the point-mass at $x \in \mathbb{R}$. Note that L_n is a random probability measure on Γ. We write

$$\mathfrak{M}_1(\Gamma) := \left\{ \nu = (\nu_1, \nu_2, \ldots, \nu_r) \in [0,1]^r : \sum_{s=1}^r \nu_s = 1 \right\}.$$

We define the total variation distance on $\mathfrak{M}_1(\Gamma)$ by

$$d(\mu, \nu) = \frac{1}{2} \sum_{s=1}^r |\mu_s - \nu_s|, \quad \mu, \nu \in \mathfrak{M}_1(\Gamma).$$

We note that $(\mathfrak{M}_1(\Gamma), d)$ is a Polish space.

Theorem (Sanov's Theorem for the empirical measure)
We define

$$P_n(S) := \mathbb{P}(L_n \in S), \quad S \subset X \text{ Borel}.$$

Then the sequence (P_n) satisfies the LDP with rate n and with rate function

$$I_\rho(\nu) = \sum_{s=1}^r \nu_s \log \left(\frac{\nu_s}{\rho_s} \right).$$

Theorem (Property of rate function)
(i) I_ρ is finite, continuous and strictly convex on $\mathfrak{M}_1(\Gamma)$.
(ii) $I_\rho(\nu) \geq 0$. Moreover, $I_\rho(\nu) = 0$ if and only if $\nu = \rho$.

Corollary
For all $a > 0$,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(L_n \in B_a^c(\rho)\right) = - \inf_{\nu \in B_a^c(\rho)} I_\rho(\nu),$$

where $B_a^c(\rho) := \{\nu \in \mathfrak{M}_1(\Gamma) : d(\nu, \rho) > a\}$.

Reference

[1] Frank den Hollander, Large Deviations,
American Mathematical Society, 2000,
IBSN: 0-8218-1989-5.

3 Weyl and Schrödinger systems

Definition
A *symplectic vector space* is a pair (L, A) consisting of a real vector space L and an anti-symmetric non-degenerate form A.

Definition
A *Weyl system* over (L, A) is a pair (K, W) consisting of a complex separable Hilbert space K and a continuous map $W : L \to U(K)$ satisfying the following equation:

$$W(z)W(z') = \exp\left(\frac{i}{2}A(z, z')\right) W(z + z').$$

Example
Note that $(\mathbb{C}^n, A)$, where $A(z, z') = \mathrm{Im}\langle z, z'\rangle$, is a symplectic vector space if $\mathbb{C}^n$ is regarded as a real $2n$-dimensional vector space. Then $(L^2(\mathbb{C}^n), W)$ is a Weyl system, where W is defined by the following equation: for every $z = x + iy \in \mathbb{C}^n$,

$$(W(z)f)(u) = \exp\left(-i\langle y, u\rangle - \frac{i}{2}\langle x, y\rangle\right) f(u + x).$$

This is, especially, called the *Schrödinger system*.

Theorem
The Schrödinger system is irreducible. In other words, the map W above is not decomposable into two maps to non-trivial Hilbert spaces.

4 Gaussian Hilbert spaces and Fock space

Let $(\Omega, \mathcal{F}, \mathrm{P})$ be a probability space.

Definition
A *Gaussian Hilbert space* is a closed subspace of $L^2(\Omega, \mathcal{F}, \mathrm{P})$ consisting of centred Gaussian random variables.

Let H be a Gaussian Hilbert space defined on $(\Omega, \mathcal{F}, \mathrm{P})$.

Definition
Let, for $n > 0$, $\overline{\mathcal{P}}_n(H)$ be the closure in $L^2(\Omega, \mathcal{F}, \mathrm{P})$ of the linear space $\mathcal{P}_n(H) = \{p(\xi_1, \cdots, \xi_m) : p$ is a polynomial of degree $\leq n$; $\xi_1, \cdots, \xi_m \in H$; $m < \infty\}$ and let $H^{:n:} = \overline{\mathcal{P}}_n(H) \cap \overline{\mathcal{P}}_{n-1}(H)^\perp$. For $n = 0$ we let $H^{:0:} = \overline{\mathcal{P}}_0(H)$, the space of constants.

Theorem (*Wiener chaos decomposition*)
The spaces $H^{:n:}$, $n \geq 0$, are mutually orthogonal, closed subspaces of $L^2(\Omega, \mathcal{F}, \mathrm{P})$ and

$$\bigoplus_{n=0}^{\infty} H^{:n:} = L^2(\Omega, \mathcal{F}(H), \mathrm{P}),$$

where $\mathcal{F}(H)$ is the σ-field generated by the random variables in H.

Definition
π_n denotes the orthogonal projection of $L^2(\Omega, \mathcal{F}, \mathrm{P})$ onto $H^{:n:}$. If $\xi_1, \cdots, \xi_n \in H$, their *Wick product* $:\xi_1 \cdots \xi_n: \in H^{:n:}$ is given by $:\xi_1 \cdots \xi_n: = \pi_n(\xi_1 \cdots \xi_n)$.

Let $H^{\odot n}$ be the symmetric tensor power of a Hilbert space H. Since the multiplication $(f_1 \odot \cdots \odot f_n) \odot (f_{n+1} \odot \cdots \odot f_{n+m}) = f_1 \odot \cdots \odot f_{n+m}$ may be extended to a continuous bilinear operation $H^{\odot n} \times H^{\odot m} \to H^{\odot(n+m)}$ for any $n, m \geq 0$, the direct sum $\Gamma_*(H) = \sum_{n=0}^{\infty} H^{\odot n}$ is a graded commutative algebra which is called the *symmetric tensor algebra* of H. Its completion, the Hilbert space $\Gamma(H) = \bigoplus_{n=0}^{\infty} H^{\odot n}$, is called the (symmetric) *Fock space* over H.

Now, let H be a Gaussian Hilbert space. By the properties of Wick products, we obtain the following fundamental result.

Theorem
If H is a Gaussian Hilbert space, then the map $\xi_1 \odot \cdots \odot \xi_n \mapsto :\xi_1 \cdots \xi_n:$ defines a Hilbert space isometry of $H^{\odot n}$ onto $H^{:n:}$. Taken together for all $n \geq 0$, these maps define an algebra isomorphism of the symmetric tensor algebra $\Gamma_*(H)$ onto $\bigcup_{n=0}^{\infty} \mathcal{P}_n(H)$ with the Wick multiplication; this extends to an isometry of the Fock space $\Gamma(H)$ onto $\bigoplus_{n=0}^{\infty} H^{:n:} = L^2(\Omega, \mathcal{F}(H), \mathrm{P})$.

Reference

[1] Christopher J. Fewster and Kasia Rejzner, *Algebraic Quantum Field Theory – an Introduction*, arXiv:1904.0405.

[2] John C. Baez, Irving E. Segal and Zhengfang Zhou, *Introduction to Algebraic and Constructive Quantum Field Theory*, Princeton Series in Physics. Princeton, N.J: Princeton University Press, 1992.

[3] Svante Janson, *Gaussian Hilbert Spaces*, Cambridge University Press, 2009.