Mixture of Large Deviation Systems

Jinwen Chen

Aug.10,2004

. **Motivation**–Evaluate the asymptotic values of certain Laplace integral

$$\int_{R^d} p_n(x)e^{nh(x)}dx.$$

The classical Laplace's Theorem provides explicit estimates for nice p and h. To handle more general cases, an alternative way is to apply Varadhan's Theorem in the theory of large deviations.

1 Large deviation system (LDS)

· **Large deviation principle (LDP)**

Definition. A sequence $\{P_n,\ n \geq 1\}$ on a topological space E with its Borel σ-algebra $\mathcal{B}(E)$ is said to satisfy a (full) large deviation principle with rate function I if

(1) I is lower semi-continuous on E;

(2) For every closed subset F of E,

$$\limsup_{n\to\infty} \frac{1}{n} \log P_n(F) \leq -\inf_{y\in F} I(y); \tag{1.1}$$

(3) For every open subset G of E,

$$\liminf_{n\to\infty} \frac{1}{n} \log P_n(G) \geq -\inf_{y\in G} I(y). \tag{1.2}$$

The rate function I is said to be *good* if it has compact level sets. (1.1) and (1.2) are called the large deviation upper and lower bounds respectively.

Varadhan's Theorem. *If $\{P_n,\ n \geq 1\}$ satisfies the full LDP with the rate function I, then for each $F \in C_b(E)$,*

$$\lim_{n\to\infty} \frac{1}{n} \log \int e^{nF} dP_n = \sup_y [F(y) - I(y)]. \qquad (1.3)$$

· **Large deviation systems**

Suppose that for each $x \in X$, $\{P_{n,x},\ n \geq 1\}$ is a sequence of probability measures on E that satisfies a full LDP with rate function I_x. The parameter space X is a topological space, $\mathcal{B}(X)$ is the Borel σ-algebra.

Remark 1.1. The terminology LDS was first used in [DG: Dawson and Gartner (PTRF,1994)]. In theri definition, a system $\{P_{n,x},\ n \geq 1,\ x \in X\}$ is called a LDS if for any $x \in X$ and any sequence $\{x_n,\ n \geq 1\}$ in X with $x_n \to x$, the sequence $\{P_{n,x_n},\ n \geq 1\}$ satisfies the full LDP with rate function I_x.

Earlier than this, such systems appeared in [DinZ: Dinwoodie and Zabell (AP,1992)]. They did not use the term LDS, but called a system with the above property exponentially continuous. One of the main purposes in that paper is to study large deviations for exchangeable random variables, which is a special case of mixture of LDS.

Large deviation problem for mixture of LDS

Given any probability measures π_n on X we obtain a sequence of mixture $\{P_n,\ n \geq 1\}$ defined by

$$P_n = \int P_{n,x} \pi_n(dx).$$

A natural problem is to study the LDP for $\{P_n,\ n \geq 1\}$. [DinZ] studied this problem in the case of X being compact and $\pi_n = \pi$ is independent of n. Exponential continuity plays the key role in their study. The results was then applied to exchangeable sequences.

[G: Grunwald (PTRF,1996)] studied this problem too, and applied the results obtained to the Sherrington-Kirkpatrick(SK) spin-glass models. A continuous surjection from E to X was involved.

Because of our motivating problem, one of the main purpose of our investigation on this problem is to obtain the LDP for mixture for array of some iid sequences by relaxing the conditions on the exponential continuity and on the compactness of the parameter space X. LDP for mixture of general LDS are also discussed.

2 LDP for mixture of general LDS

We assume that E is a Polish space and X is first countable. Let $\{P_{n,x},\ n \geq 1\}_{x\in X}$, $\{\pi_n,\ n \geq 1\}$ and $\{P_n,\ n \geq 1\}$ be as in Section 1. Assume that for each n, $P_{n,x} = P_n(x,\ \cdot)$ is a probability kernel on $X \times E$. In the general cases, our main result is the following

Theorem 2.1. *Suppose that* $\{P_{n,x},\ n \geq 1\}_{x\in X}$ *is exponentially continuous with the family of rate function* $\{I_x,\ x \in X\}$.

(1) If $\{\pi_n,\ n \geq 1\}$ *has the large deviation upper bounds (1.1) (resp. lower bounds (1.2)) with the good rate function* $\bar{J}$ *(resp.* $\underline{J}$*), then* $\{P_n,\ n \geq 1\}$ *has the large deviation upper bounds (resp. lower boundes) with a rate function* $\bar{I}$ *(resp.* $\underline{I}$*) that is the lower semi-continuous version of* $\inf_x[I_x + \bar{J}(x)]$ *(resp.* $\inf_x[I_x + \underline{J}(x)]$*);*

(2) If $\{\pi_n,\ n \geq 1\}$ *satisfies the full LDP with the good rate function* J*, then* $\{P_n,\ n \geq 1\}$ *satisfies the full LDP with the good rate function* I *which is the lower*

semi-continuous version of $\inf_x[I_x + J(x)]$.

Remark 2.1. If X is compact and $\pi_n = \pi$, $\forall n \geq 1$, then $J \equiv 0$ and thus $I = \inf_x I_x$. If in addition $I_x(y)$ is jointly lower semi-continuous in $(x,\ y)$, then I is also lower semi-continuous.

3 LDP for mixture of array of iid sequences

Let X, E and $\{\pi_n,\ n \geq 1\}$ be as before, $\{P_x,\ x \in X\}$ a family of probability measures on some measurable space $(\Omega, \mathcal{F})$ and for each $n \geq 1$, $\{X_k^n,\ k \geq 1\}$ be sequence of E-valued random variables that is iid under each P_x. Define

$$Y_n = \frac{1}{n}\sum_{k=1}^{n} X_k^n$$

and let $P_{n,x} = P(Y_n \in \cdot)$ be the law of Y_n under P_x. Then define

$$P_n = \int P_{n,x}\pi_n(dx).$$

A direct consequence of Theorem 2.1 is the following

Corollary 3.1. If $\{P_{n,x},\ n \geq 1\}_{x\in X}$ is exponentially continuous with the family of rate functions $\{I_x,\ x \in X\}$ and $\{\pi_n,\ n \geq 1\}$ satisfies the full LDP with the good rate function J, then $\{P_n,\ n \geq 1\}$ satisfies the full LDP with the rate function I that is the lower semi-continuous version of $\inf_x[I_x + J(x)]$.

To relax the conditions on the exponential continuity in the above theorem and on the tightness of X in [DinZ], we consider the special case where $\pi_n = \pi$ is independent of n. Instead, we impose certain moment conditions. We also need the exponential tightness of $\{P_n,\ n \geq 1\}$.

Definition. $\{P_n,\ n \geq 1\}$ is said to be exponentially tight, if for every $L > 0$, there is a compact subset K_L

of E such that

$$P_n(K_L^c) \le e^{-nl}, \quad \forall n \ge 1. \tag{3.1}$$

Remark 3.1. It is well known that the exponential tightness is necessary for obtaining the LDP with a *good* rate function.

The following are our main results.

Theorem 3.2. *Under the above notations, assume that $\{P_n,\ n \ge 1\}$ is exponentially tight. If for each $\lambda \in E'$, the topological dual of E,*

$$M_{n,x}(\lambda) = \int e^{<\lambda, X_1^n>} dP_x < \infty,$$

and $M_{n,x}(\lambda)$ converges uniformly in x as $n \to \infty$ to some functional $M_x(\lambda)$ for which $\frac{d}{dt} M_x(t\lambda_1 + (1-t)\lambda_2)|_{t=0}$ exists for any λ_1 and λ_2, then $\{P_n,\ n \ge 1\}$ satisfies the full LDP with a good rate function.

In particular, if $\{X_k^n\} = \{X_k\}$ is independent of n and the moment generating function $M_x(\lambda)$ of X_1 w.r.t. P_x is finite on E' for each x, then the above conclusion holds.

Remark 3.2. (1) The technique for proving the full LDP is to apply the Inverse Varadhan Theorem.

(2) the moment condition can be weakened.

(3) If X is compact and $\{X_k^n\}$ and $\pi_n = \pi$ are independent of n then $I = \inf_x I_x$, where I_x is the good rate function for LDP of $\{P_x(Y_n \in \cdot),\ n \ge 1\}$. But if X is not compact, the above equality need not be true.

Remark 3.3. An important case is that where $X_k = \delta_{Z_k}$ is the Dirac measure concentrated on $\{Z_k\}$ for some E-valued sequence $\{Z_k,\ k \ge 1\}$ which is iid under each P_x. In this case Y_n's are the associated empirical measures. It is obvious that the moment condition is satisfied. Thus to obtain the full LDP, we only need to seek for the required exponential tightness. The following is a prctical sufficient condition:

Theorem 3.3. *If there is a $\delta > 0$ and a measurable function h on E which is bounded from below and has compact level sets, such that*

$$\limsup_{n\to\infty} \frac{1}{n} \log \int \pi(dx)[\int e^{\delta h(Z_1)} dP_x]^n < \infty,$$

then $\{P_n = P(Y_n \in \cdot),\ n \geq 1\}$ is exponentially tight.

4 Discussion on potential applications

1. To Laplace integrals.

Consider the following very simple case. Let $-\infty < a < b < \infty$, to estimate

$$r_n = \int_a^b e^{nh(x)} \pi_n(dx).$$

Suppose that $-\infty < \underline{h} = \inf h \leq \sup h = \overline{h} < \infty$ and that $\{\pi_n,\ n \geq 1\}$ satisfies the full LDP with the good rate function J. Then it cane be checked that

$$\lim_{n\to\infty} \frac{1}{n} \log r_n = \sup_x [h(x) + J(x)].$$

2. To certain "hidden" Markov models

Let $\{X_n,\ n \geq 1\}$ be a "strongly ergodic" Markov chain on some probability space $(\Omega,\ \mathcal{F},\ \mathcal{P})$ with the state space $E \subset \mathbf{R}$, ν its unique invariant probability measure, $\{Y_n,\ n \geq 1\}$ be iid under P with the distribution $P(X_1 = 0) = 1 - p = 1 - P(X_1 = 1)$. The two families are assumed to be independent of each other. Define $Z_n = X_n Y_n$ (a Markove chain with deletion?). $Y_n = \frac{1}{n} \sum_{k=1}^n Z_k$ and $P_n = P(Y_n \in \cdot)$. Under suitable conditions it can be varified that $\{P_n,\ n \geq 1\}$ satisfies the full LDP with the rate function

$$I(y) = \sup_{\lambda}[\lambda y - \int \Lambda(\lambda x)\nu(dx)].$$

Other possible applications....

The end!

Thank you!

References