Contraction Principles some applications

Fabrice Gamboa (Institut de Mathématiques de Toulouse)

7th of June 2017

Christian and Patrick 59th Birthday

Overview

Overview

Starter I : Large deviations

$(\mathbb{P}_n)$ on E satisfies a LDP with good rate I means that at log scale

$$\forall A \text{ measurable set, } \mathbb{P}_n(A) \approx \exp(-n \inf_{x \in A} I(A)).$$

with

- $I \geqslant 0$ is lower semi-continuous,
- $\forall L > 0, \{x \in E : \ I(x) \leqslant L\}$ is compact.
- I is unique

Starter II : Contraction principles

Contraction principle

$f : E \to F$ a continuous map

$(\mathbb{P}_n)$ on E LDP with good rate $I \Rightarrow (\mathbb{P}_n \circ f^{-1})$ LDP on F with good rate

$$J(\cdot) := \inf_{\{x : f(x) = \cdot\}} I(x).$$

Inverse contraction principle

$g : F \to E$ a continuous bijective map + exponential tightness of $(\mathbb{P}_n)$

$(\mathbb{P}_n \circ g^{-1})$ on E LDP with rate $I \Rightarrow (\mathbb{P}_n)$ LDP on F with good rate

$$J(\cdot) := I(g(\cdot)).$$

Starter III : Kullback Leibler Divergence I

R. Leibler (1914-2003)

S. Kullback (1907-1994)

P, Q probability measures on some space E

$$K(P,Q) = \begin{cases} \int_E \log \frac{dP}{dQ} dP & \text{if } P \ll Q \text{ and } \log \frac{dP}{dQ} \in L^1(P) \\ +\infty & \text{otherwise} \end{cases}$$

Starter III : large deviations two examples

- Sanov Theorem : (X_n) i.i.d. Q on E

  $\Rightarrow F_n(dx) = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}(dx)$ LDP on $\mathbb{P}(E)$ rate $K(\cdot, Q)$

- Weighted measures : (Z_n) i.i.d. $\mathcal{E}(1)$, $(x_n) \in$ supp $P \subset E$, $\frac{1}{n}\sum_{i=1}^{n} \delta_{x_i}(dx) \to P$

  $\Rightarrow \nu_n(dx) = \frac{1}{n}\sum_{i=1}^{n} Z_i\delta_{x_i}(dx)$ LDP on $\mathbb{M}_+(E)$ rate $K(P, \cdot)$

Overview

1. Appetizer : Christian and Patrick secret lives
2. Starter : Around large deviations
3. Main course : Sanov and Markov
4. Trou normand : Marginal problem and LDP
5. Cheese : Markov generalized moment problem and LDP
6. Dessert : Sum rules

Main course : Sanov and Markov

C. Léonard

P. Cattiaux

- Cattiaux, P., & Léonard C. (1994). Minimization of the Kullback information of diffusion processes. Annales de l'IHP probabilités et statistique, 30, 83-132.
- Cattiaux, P., & Léonard C. (1995). Large deviations and Nelson processes. Forum Mathematicum, 7, 95-116.
- Cattiaux, P., & Léonard C. Minimization of the Kullback information for some Markov processes. (1996). Séminaire de Probabilités *XXX*, Springer, 288-311

Sanov Theorem : Q law of a Markov process on $[0, T]$, (X_n) i.i.d. Q

$$F_n(dx) = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}(dx)$$

Random law of a random process-Satisfies LDP rate $K(\cdot, Q)$

Main course (bis) : Sanov and Markov

C. Léonard

P. Cattiaux

Various computation of the contracted functional

$$J(\cdot) := \inf_{\{x : f(P) = \cdot\}} K(P, Q).$$

$f(P)$ is the marginals flow of P

Feasibility conditions for the optimization problem (existence of solution having finite entropy)

Overview

Trou normand : Marginal problem and LDP

P. Cattiaux

F.G

- Cattiaux, P., & G F. (1999). Large deviations and variational theorems for marginal problems. Bernoulli, 5, 81-108.

P on $E \times F$, ν on E, μ on F données

$\exists\, Q \ll P$ on $E \times F$ marginals ν and μ+conditions ?

Tool : LDP weighted measures (ν_n) +contraction

- (Z_n) i.i.d., $(x_n, y_n) \in E \times F$,
- $\frac{1}{n} \sum_{i=1}^{n} \delta_{(x_i, y_i)}(dx, dy) \to P$

$$\nu_n(dx, dy) = \frac{1}{n} \sum_{i=1}^{n} Z_i \delta_{(x_i, y_i)}(dx, dy).$$

Trou normand (bis) : Marginal problem and LDP

P. Cattiaux

F.G

$$\nu_n(dx, dy) = \frac{1}{n}\sum_{i=1}^{n} Z_i \delta_{(x_i,y_i)}(dx, dy).$$

(ν_n) satisfies a LDP with good rate I

$$I(Q) := \int_{E\times F} \psi^*\left(\frac{dQ}{dP}\right) dP, \; Q \ll P.$$

ψ^* is the Cramér transform of Z

Trou normand (ter) : Marginal problem and LDP

P. Cattiaux

F.G

Contraction on the two marginals

Example : $Z \sim L\mathcal{B}(1/2)$ $(L > 1)$

$$\exists\, Q \ll P \text{ on } E \times F \text{ having marginals } \nu \text{ and } \mu \text{ with } \frac{dQ}{dP} \leqslant L$$

$$\text{iff } \forall h_1 \in C(E), h_2 \in C(F),$$

$$\int_E h_1 d\nu + \int_F h_2 d\mu \leqslant \int_{E\times F} \log\left(1 + \exp\left(L(h_1(x) + h_2(y))\right)\right) P(dx, dy)$$

Overview

Cheese : Markov generalized moment problem and LDP

D. Dacunha-Castelle

F.G

- Dacunha-Castelle, D., & G F. (1990). Maximum d'entropie et problème des moments. Annales de l'IHP probabilités et statistique, 26, 567-596.

$P = dx$ on [0,1] and contraction on the Φ-moments ($\Phi \in C^k([0,1])$ given)

The feasibility of the Markov moment problem for $c \in \mathbb{R}^k$

$$\exists G \in \mathbb{M}_+([0,1])\ G \ll dx \text{ on } [0,1] \text{ with } \int_0^1 \Phi(x)G(dx) = c \text{ and } \frac{dG}{dx} \leqslant 1$$

$$\text{iff } \forall v \in \mathbb{R}^k, \quad \langle v, c\rangle \leqslant \int_0^1 \log\left(1 + \exp\langle v, \Phi(x)\rangle\right) dx$$

Overview

Sum rules and LDP coworks with

J. Nagel (Munich and EURANDOM)

A. Rouault (Versailles)

Orthogonal polynomial recursion on $\mathbb{T}$

- μ probability measure on $\mathbb{T}$
  1) (p_n) sequence of orthogonal polynomials associated to μ
  2) p_n is monic and has degree n
  3) $k \neq n$, $\int_{\mathbb{T}} p_n(z) p_k(z) \mu(dz) = 0$
- Satisfies the recursion

  $\rightarrow$ $p_{n+1}(z) = zp_n(z) - \overline{\alpha}_n p_n^*(z)$ where $p_n^*(z) := z^n \overline{p_n(1/\bar{z})}$.
  $\rightarrow$ $\alpha_n = -p_{n+1}(0)$ is the Verblunsky coefficient

Orthogonal polynomial recursion on $\mathbb{T}$

- μ probability measure on $\mathbb{T}$
  1) (p_n) sequence of orthogonal polynomials associated to μ
  2) p_n is monic and has degree n
  3) $k \neq n, \int_{\mathbb{T}} p_n(z) p_k(z) \mu(dz) = 0$
- Satisfies the recursion

$\rightarrow$ $p_{n+1}(z) = zp_n(z) - \overline{\alpha}_n p_n^*(z)$ where $p_n^*(z) := z^n \overline{p_n(1/\bar{z})}$.
$\rightarrow$ $\alpha_n = -p_{n+1}(0)$ is the Verblunsky coefficient

Orthogonal polynomial recursion on $\mathbb{T}$

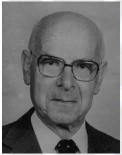
- μ probability measure on $\mathbb{T}$
  1) (p_n) sequence of orthogonal polynomials associated to μ
  2) p_n is monic and has degree n
  3) $k \neq n, \int_{\mathbb{T}} p_n(z) p_k(z) \mu(dz) = 0$
- Satisfies the recursion

  $\rightarrow$ $p_{n+1}(z) = z p_n(z) - \overline{\alpha}_n p_n^*(z)$ where $p_n^*(z) := z^n \overline{p_n(1/\bar{z})}$.

  $\rightarrow$ $\alpha_n = -p_{n+1}(0)$ is the Verblunsky coefficient

Verblunsky version of Szegö Theorem

G. Szegö (1895-1985)



S. Verblunsky (1906-1996)

Verblunsky-Szegö Theorem

λ Lebesgue measure on $\mathbb{T}$.

$$K(\lambda, \mu) = -\sum_{n=0}^{\infty} \log(1 - |\alpha_n|^2)$$

Second ingredient : the three parametrisations

- μ probability measure on $\mathbb{R}$
  - $\rightarrow$ Assume that μ has all its **moments finite**
  - $\rightarrow$ (p_n) normalized othogonal polynomials in $L^2(\mu)$
  - $\rightarrow$ Three terms recursion

$$xp_n(x) = a_n p_{n-1}(x) + b_{n+1} p_n(x) + a_{n+1} p_{n+1}(x), \quad a_n > 0, \; b_{n+1} \in \mathbb{R}$$

- Assume moreover that μ is supported on $[0, +\infty[$
  $b_n = z_{2n-2} + z_{2n-1}$, and $a_n^2 = z_{2n-1} z_{2n}$.
- If μ is supported on $[0, 1]$, μ may be seen as the pushforward of a measure on $\mathbb{T}$ invariant by $2\pi - \theta$ by $\sin^2(\theta/2)$.
  $b_{k+1} = 1/4[2 - (1 - \alpha_{2k-1})\alpha_{2k} - (1 + \alpha_{2k-1})\alpha_{2k-2}]$ and
  $a_{k+1} = 1/4\sqrt{(1 - \alpha_{2k-1})(1 - \alpha_{2k}^2)(1 + \alpha_{2k+1})}$

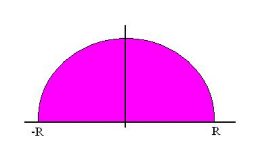
Second ingredient : the three parametrisations

- μ probability measure on $\mathbb{R}$
  - $\rightarrow$ Assume that μ has all its **moments finite**
  - $\rightarrow$ (p_n) normalized othogonal polynomials in $L^2(\mu)$
  - $\rightarrow$ Three terms recursion

$$xp_n(x) = a_n p_{n-1}(x) + b_{n+1} p_n(x) + a_{n+1} p_{n+1}(x), \quad a_n > 0, \; b_{n+1} \in \mathbb{R}$$

- Assume moreover that μ is supported on $[0, +\infty[$
  $b_n = z_{2n-2} + z_{2n-1}$, and $a_n^2 = z_{2n-1} z_{2n}$.
- If μ is supported on $[0, 1]$, μ may be seen as the pushforward of a measure on $\mathbb{T}$ invariant by $2\pi - \theta$ by $\sin^2(\theta/2)$.
  $b_{k+1} = 1/4[2 - (1 - \alpha_{2k-1})\alpha_{2k} - (1 + \alpha_{2k-1})\alpha_{2k-2}]$ and
  $a_{k+1} = 1/4\sqrt{(1 - \alpha_{2k-1})(1 - \alpha_{2k}^2)(1 + \alpha_{2k+1})}$

Semicircular distribution



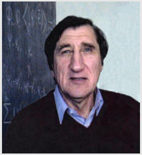
$$R = 2$$

$$SC(dx) = \frac{1}{2\pi}\sqrt{4 - x^2}\,\mathbb{1}_{[-2,2]}(x)\,dx$$

Limit of the eigenvalues distribution of the GUE

$$a_k = 1,\ b_k = 0 \text{ for all } k \geqslant 1\,.$$

Pastur-Marchenko Distribution



L. Pastur

V. Marchenko

$$\mathrm{MP}_\tau(dx) = \frac{\sqrt{(\tau^+ - x)(x - \tau^-)}}{2\pi\tau x} \mathbb{1}_{(\tau^-,\tau^+)}(x)dx \ , \tau^\pm = (1 \pm \sqrt{\tau})^2$$

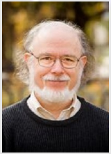
Limit of the squared singular values distribution of rectangular Gaussian matrices. $\tau \in]0, 1]$ the asymptotic ratio nb col/ nb line

$$a_k = \sqrt{\tau}\ (k \geqslant 1)\ ,\ \ b_1 = 1\ ,\ b_k = 1 + \tau\ \ (k \geqslant 2)$$

and correspond to $z_{2n-1} = 1$ and $z_{2n} = \tau$ for all $n \geqslant 1$.

Kesten Mac Kay Distribution

H. Kesten



B. Mc Kay

$$\mathrm{KMK}_{\kappa_1,\kappa_2}(dx) = \frac{(2+\kappa_1+\kappa_2)}{2\pi} \frac{\sqrt{(u^+-x)(x-u^-)}}{x(1-x)} \mathbb{1}_{(u^-,u^+)}(x)dx$$

$$u^\pm := \frac{1}{2} + \frac{\kappa_1^2 - \kappa_2^2 \pm 4\sqrt{(1+\kappa_1)(1+\kappa_2)(1+\kappa_1+\kappa_2)}}{2(2+\kappa_1+\kappa_2)^2}, \kappa_1, \kappa_2 \geqslant 0.$$

Asymptotic distribution of the eigenvalues in the Jacobi-ensemble

The associated Verblunsky coefficients for $k \geqslant 0$,

$$\alpha_{2k} = \frac{\kappa_1 - \kappa_2}{2 + \kappa_1 + \kappa_2}, \quad \alpha_{2k+1} = -\frac{\kappa_1 + \kappa_2}{2 + \kappa_1 + \kappa_2}.$$

Then

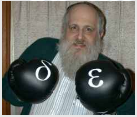
$$a_1 = \frac{\sqrt{(1+\kappa_1)(1+\kappa_2)}}{(2+\kappa_1+\kappa_2)^{3/2}}, \quad b_1 = \frac{1+\kappa_2}{2+\kappa_1+\kappa_2},$$

and for $k \geqslant 2$

$$a_k = \frac{\sqrt{(1+\kappa_1+\kappa_2)(1+\kappa_1)(1+\kappa_2)}}{(2+\kappa_1+\kappa_2)^2}, \quad b_k = \frac{1}{2}\left[1 - \frac{\kappa_1^2 - \kappa_2^2}{(2+\kappa_1+\kappa_2)^2}\right].$$

Killip Simon Theorem

R. Killip



B. Simon

$$K(SC|\mu) + \sum_{n=1}^{N^+} \mathcal{F}_H^+(\lambda_n^+) + \sum_{n=1}^{N^-} \mathcal{F}_H^-(\lambda_n^-) = \sum_{k\geqslant 1} \left(\frac{1}{2} b_k^2 + a_k^2 - 1 - \log(a_k^2)\right)$$

$$\mathcal{F}_H^+(x) := \begin{cases} \int_2^x \sqrt{t^2-4}\,dt = \frac{x}{2}\sqrt{x^2-4} - 2\log\left(\frac{x+\sqrt{x^2-4}}{2}\right) & \text{if } x \geqslant 2 \\ \infty \quad \text{otherwise,} & \end{cases}$$

$$\mathcal{F}_H^-(x) := \mathcal{F}_H^+(-x)$$

Our sum rules I

$$K(MP_\tau|\mu) + \sum_{n=1}^{N^+} \mathcal{F}_L^+(\lambda_n^+) + \sum_{n=1}^{N^-} \mathcal{F}_L^-(\lambda_n^-) = \sum_{k=1}^{\infty} \tau^{-1} G(z_{2k-1}) + G(\tau^{-1} z_{2k})$$

$$\mathcal{F}_L^+(x) = \begin{cases} \int_{\tau^+}^{x} \frac{\sqrt{(t-\tau^-)(t-\tau^+)}}{t\tau}\, dt & \text{if } x \geqslant \tau^+, \\ \infty & \text{otherwise,} \end{cases}$$

$$\mathcal{F}_L^-(x) = \begin{cases} \int_{x}^{\tau^-} \frac{\sqrt{(\tau^- - t)(\tau^+ - t)}}{t\tau}\, dt & \text{if } x \leqslant \tau^-, \\ \infty & \text{otherwise.} \end{cases}$$

$$G(x) = x - 1 - \log x, \ (x > 0).$$

Our sum rules II

$$K(KMK_{\kappa_1,\kappa_2}|\mu) + \sum_{n=1}^{N^+} \mathcal{F}_J^+(\lambda_n^+) + \sum_{n=1}^{N^-} \mathcal{F}_J^-(\lambda_n^-) = \sum_{k=0}^{\infty} H_1(\alpha_{2k+1}) + H_2(\alpha_{2k})$$

$$\mathcal{F}_J^+(x) = \begin{cases} \int_{u^+}^{x} \frac{\sqrt{(t-u^+)(t-u^-)}}{t(1-t)}\,dt & \text{if } u^+ \leqslant x \leqslant 1 \\ \infty & \text{otherwise.} \end{cases}$$

$$\mathcal{F}_J^-(x) = \begin{cases} \int_{x}^{u^-} \frac{\sqrt{(u^- - t)(u^+ - t)}}{t(1-t)}\,dt & \text{if } 0 \leqslant x \leqslant u^- \\ \infty & \text{otherwise.} \end{cases}$$

For $-1 \leqslant x \leqslant 1$

$$H_1(x) = -(1+\kappa_1+\kappa_2)\log\left[\frac{2+\kappa_1+\kappa_2}{2(1+\kappa_1+\kappa_2)}(1-x)\right] - \log\left[\frac{2+\kappa_1+\kappa_2}{2}(1+x)\right]$$

$$H_2(x) = -(1+\kappa_1)\log\left[\frac{(2+\kappa_1+\kappa_2)}{2(1+\kappa_1)}(1+x)\right] - (1+\kappa_2)\log\left[\frac{(2+\kappa_1+\kappa_2)}{2(1+\kappa_1)}(1-x)\right].$$

Particular case $\kappa_1 = \kappa_2 = 0 \Rightarrow u^- = 0,\ u^+ = 1$, KMK_{κ_1,κ_2} arsine law.

We recover the Szegö Theorem pushed on $[0, 1]$ by $\sin^2(\theta/2)$.

How it works ?

Main idea

Large deviations +inverse contraction principle

Random measures

Spectral random measures associated to classical ensembles of random matrices (A_n)

(ν_n) defined by its moments

$$\int x^k \nu_n(dx) := \langle e, A_n^k e\rangle, \ (k \in \mathbb{N}), \ \|e\| = 1, e \text{ cyclic}$$

Main tool I

Theorem

- ν_n is supported by the eigenvalues of A_n.
- For the classical ensembles the weights and the supporting points are independents
- Furthermore, in this last case, the vector of weights is Dirichlet distributed.

Main tool II

Theorem

1) In the GUE, (a_k^n) and (b_k^n) are independent sequences of independent variables (normal and χ^2)
2) In the Wishart model, (z_k^n) is a sequence of independent variables (χ^2).
3) In the Jacobi model, (α_k^n) is a sequence of independent variables (Beta)

Some references around the dessert

- Verblunsky, S. (1936). On positive harmonic functions. Proceedings of the London Mathematical Society, 2(1), 290-320.
- Killip, R., & Simon, B. (2003). Sum rules for Jacobi matrices and their applications to spectral theory. Annals of mathematics, 253-321.
- Simon, B. (2010). Szegö's theorem and its descendants : spectral theory for L2 perturbations of orthogonal polynomials. Princeton University Press.
- GF., Nagel, J., & Rouault, A. (2016). Sum rules via large deviations. Journal of Functional Analysis, 270(2), 509-559.
- GF., Nagel, J., & Rouault, A. (2016). Sum rules and large deviations for spectral matrix measures. arXiv preprint arXiv :1601.08135.
- GF., Nagel, J., & Rouault, A. (2017). Sum rules and large deviations for spectral measures on the unit circle. Random Matrices : Theory and Applications.
- Breuer, J., Simon, B., & Zeitouni, O. (2017). Large Deviations and Sum Rules for Spectral Theory-A Pedagogical Approach. To appear in Jounal of Spectral Theory.
- Breuer, J., Simon, B., & Zeitouni, O. (2017). Large deviations and the Lukic conjecture. arXiv preprint arXiv :1703.00653.