Moderate deviations for random matrices

Fuqing Gao

School of Mathematics and Statistics, Wuhan University

8th Workshop on Markov Processes and Ralated Topics
July 16 to 21, 2012, in Beijing and Wuyi Mouuntain Villa

This talk is based on joint works with Lei Chen and Shaochen Wang

Outline

1. Wigner matrix and Wishart process
   - Wigner matrix
   - Wishart process
2. Cramér type moderate deviations for eigenvalues of Wigner matrices
   - Cramér type moderate deviations
   - Sketch of the proof
   - Berry-Esseen bounds
   - Cramér type moderate deviations for covariance matrices
3. Moderate deviations for small perturbation Wishart processes
   - Small perturbation Wishart processes
   - Moderate deviations for small perturbation Wishart processes
   - Moderate deviations for the eigenvalue processes
   - Sketch of the proof

Outline

- A Wigner Hermitian matrix: $W_n = M_n/\sqrt{n}$, where $M_n = (m_{ij})_{n\times n}$ is a random Hermitian $n \times n$ matrix such that
  - $\{\Re m_{ij}, \Im m_{ij} : 1 \le i < j \le n\}$ is a collection of i.i.d. real random variables (with mean zero and variance $1/2$.), where $\Re m_{ij}$ and $\Im m_{ij}$ denotes real part and image part of m_{ij}, respectively,
  - $m_{ii}, 1 \le i \le n$ are i.i.d. real random variables (with mean zero and variance 1).
- Gaussian unitary ensemble (GUE): A Wigner Hermitian matrix which the entries are Gaussian random variables.
- The empirical distribution function of the eigenvalues,

$$F_n := \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i},$$

where $\lambda_1 < \cdots \lambda_n$ are the eigenvalues of W_n.

- The semicircle law (Wigner(1955, 1958)),

$$F_n(x) \to d\rho_{sc}(x) = \frac{1}{2\pi}\sqrt{4-x^2}I_{[-2,2]}(x)dx.$$

- The large deviations and the moderate deviations for the empirical distributions of the eigenvalues (Ben Arous, Guionnet (1997), Dembo, Guionnet, Zeitouni(2003)).
- The central limit theorem and The rough moderate deviations for the eigenvalue of Wigner matrices in the bulk and at the edge cases (Gustavsson(2005),O'Rourke (2010),Su (2006), Döring, Eichelsbacher(2010)).
- This talk introduces the Berry-Esseen bounds and the Cramér type moderate deviations for the eigenvalue of Wigner matrices in the bulk and at the edge cases.

Outline

- $\mathcal{S}_m^+$ is the set of $m \times m$ real symmetric non-negative matrices with the Hilbert-Smith norm $\|\cdot\|_{HS}$.
- $\widetilde{\mathcal{S}}_m^+$ denotes the set of $m \times m$ positive definite symmetric matrices.
- Wishart matrix: $W = X^T X$, where X is a Gaussian random matrix.
- Wishart process: The solution of the following SDE with matrix values:

$$\begin{cases} dX_t = \sqrt{X_t}dB_t + dB_t^T\sqrt{X_t} + \rho I_m dt; \\ X_0 = \mathbf{x}, \end{cases} \tag{1.1}$$

  where $\{B_t, t \geq 0\}$ is a $m \times m$ matrix valued Brownian motion, B^T denotes the transpose of the matrix B and I_m is the identity matrix.
- It has a unique strong solution in $\widetilde{\mathcal{S}}_m^+$ under the condition $\rho \geq m+1$ and $\mathbf{x} \in \widetilde{\mathcal{S}}_m^+$ (Bru (1991)).
- Its distribution on $C(\mathbb{R}_+, \mathcal{S}_m^+)$ is denoted by $\mathbf{Q}_\mathbf{x}^\rho$.

- Let $\lambda_1(t), \cdots, \lambda_m(t)$, $t \in [0, T]$ be the ordered eigenvalue processes of X, then $\lambda_1(t) < \cdots < \lambda_m(t)$, $t \in [0, T]$, and it is the solution of the following SDE (Bru (1991))

$$\begin{cases} d\lambda_i(t) = 2\sqrt{\lambda_i(t)}d\beta_i(t) + \left(\rho \sum_{k \neq i} \frac{\lambda_i(t)+\lambda_k(t)}{\lambda_i(t)-\lambda_k(t)}\right) dt, \quad i = 1, \cdots, m; \\ \lambda_i(0) = \lambda_i, \end{cases}$$

  where β_i, $i = 1, \cdots, m$ are independent Brownian motions.
- The large deviations for the empirical distributions of the eigenvalue of the Wishart matrices (Hiai, Petz (1998)).
- Donati-Martin (2008) considered large deviations for small perturbation Wishart processes and the eigenvalue processes.
- This talk introduces moderate deviations for small perturbation Wishar processes and the eigenvalue processes.

Let W_n be a Wigner Hermitian matrix with ordered eigenvalues $\lambda_1 \leq \cdots \leq \lambda_n$. Let $F_k(\cdot)$ denote the distribution function of Λ_k, where

$$\Lambda_k := \frac{\lambda_k - t\sqrt{2n}}{\left(\frac{\log n}{4(1-t^2)n}\right)^{1/2}}, \qquad \Lambda_{n-k} := \frac{\lambda_{n-k} - \sqrt{2n}\left(1 - \left(\frac{3\pi k}{4\sqrt{2}n}\right)^{2/3}\right)}{\left(\left(\frac{1}{12\pi}\right)^{2/3} \frac{2\log k}{2n^{1/3}k^{2/3}}\right)^{1/2}},$$

with $t = t(k,n) = G^{-1}(k/n)$ and

$$G(t) = \frac{2}{\pi}\int_{-1}^{t}\sqrt{1-x^2}dx, \quad -1 \leq t \leq 1. \tag{2.1}$$

We additionally assume the distributions of Wigner Hermitian matrix have exponential decay, i.e., there are constants C, $C' > 0$ such that

$$P(|m_{ij}| \geq t^C) \leq e^{-t} \text{ for all } t \geq C'. \tag{2.2}$$

Outline

Theorem 2.1 (Cramér type moderate deviations)

Let ρ be any fixed constant.

(1). *If λ_k is in the bulk, i.e., $k = k(n)$ such that $k/n \to a \in (0,1)$ as $n \to \infty$, then on the interval* $[0, \rho(\log n)^{1/6}]$,

$$\frac{1 - F_k(x)}{1 - \Phi(x)} = 1 + O(1)\frac{1 + x^3}{\sqrt{\log n}}, \quad \frac{F_k(-x)}{\Phi(-x)} = 1 + O(1)\frac{1 + x^3}{\sqrt{\log n}}.$$

(2). *If λ_{n-k} is at the edge, i.e., k such that $k \to \infty$ and $k/n \to 0$ as $n \to \infty$, and $\lim_{n\to\infty} \frac{k^{5/3}}{n^{2/3}\left(\log k\right)^{1/2}} = 0$, then on the interval* $[0, \rho(\log k)^{1/6}]$,

$$\frac{1 - F_{n-k}(x)}{1 - \Phi(x)} = 1 + O(1)\frac{1 + x^3}{\sqrt{\log k}}, \quad \frac{F_{n-k}(x)}{1 - \Phi(x)} = 1 + O(1)\frac{1 + x^3}{\sqrt{\log k}}.$$

Outline

1. Wigner matrix and Wishart process
   - Wigner matrix
   - Wishart process
2. Cramér type moderate deviations for eigenvalues of Wigner matrices
   - Cramér type moderate deviations
   - Sketch of the proof
   - Berry-Esseen bounds
   - Cramér type moderate deviations for covariance matrices
3. Moderate deviations for small perturbation Wishart processes
   - Small perturbation Wishart processes
   - Moderate deviations for small perturbation Wishart processes
   - Moderate deviations for the eigenvalue processes
   - Sketch of the proof

- By virtue of Tao and Vu's Four Moment Theorems (cf.[22]), Theorem 2.1 can be deduced to the Gaussian unitary ensembles (GUE).
- Since the counting function of eigenvalues of the GUE matrices forms a determinantal point process (cf.[1]) and has same distribution with an infinite sum of independent Bernoulli random variables, thus the problems are deduced to study the corresponding problems for infinite sum of independent Bernoulli random variables.
- Because the asymptotic properties of Theorem 2.1 hold uniformly in some unbounded intervals, we need some fine analysis and estimates for mean and variance of the counting function of eigenvalues.

Outline

Theorem 2.2 (Berry-Esseen bounds)

There exists a constant C such that

(1). *Under the condition of Theorem 2.2(1),*

$$\sup_{x\in\mathbb{R}} |F_k(x) - \Phi(x)| \leq C(\log n)^{-1/2},$$

(2). *Under the condition of Theorem 2.2(2),*

$$\sup_{x\in\mathbb{R}} |F_{n-k}(x) - \Phi(x)| \leq C(\log k)^{-1/2}.$$

Outline

Following exactly the same scheme as Wigner matrices, we can establish the Berry-Essen bounds and the Cramér type moderate deviations for eigenvalues of covariance matrices.

Outline

Let us consider the following $\widetilde{\mathcal{S}}_m^+$ valued diffusion with small parameter

$$\begin{cases} dX_t^\varepsilon = \sqrt{\varepsilon}(\sqrt{X_t^\varepsilon}dB_t + dB_t^T\sqrt{X_t^\varepsilon}) + \rho I_m dt, & t \le T; \\ X_0^\varepsilon = \mathbf{x}, \end{cases} \tag{3.1}$$

where $\rho > 0$.
As ε goes to 0, the solution of (3.1), X_t^ε, should converge to X^0 which is determined by the following ordinary differential equation:

$$\begin{cases} dX_t^0 = \rho I_m dt, & t \le T; \\ X_0^0 = \mathbf{x}. \end{cases} \tag{3.2}$$

Our problems:

- The asymptotic behavior of

$$Z_t^\varepsilon := \frac{X_t^\varepsilon - X_t^0}{h(\varepsilon)\sqrt{\varepsilon}},$$

where $1 \leq h(\varepsilon) = o(1/\sqrt{\varepsilon})$. When $h(\varepsilon) = 1$, this is central limit theorem; when

$$h(\varepsilon) \to +\infty \quad \text{and} \quad \sqrt{\varepsilon}h(\varepsilon) \to 0, \quad \text{as } \varepsilon \to 0, \tag{3.3}$$

this is a moderate deviation problem.

- The moderate deviations for the eigenvalue process λ_t^ε of X^ε:

$$\begin{cases} d\lambda_i^\varepsilon(t) = 2\sqrt{\varepsilon\lambda_i^\varepsilon(t)}d\beta_i(t) + \left(\rho + \varepsilon\sum_{k\neq i}\frac{\lambda_i^\varepsilon(t)+\lambda_k^\varepsilon(t)}{\lambda_i^\varepsilon(t)-\lambda_k^\varepsilon(t)}\right)dt, \quad i = 1,\cdots, \\ \lambda_i^\varepsilon(0) = \lambda_i, \end{cases}$$

- Ma, Wang and Wu ([13]) studied the moderate deviations for stochastic differential equations with Lipschtiz coefficients.
- In the case of the Wishart process, the diffusion coefficient (3.1) is not Lipschtiz continuous.
- In order to study moderate deviations for the eigenvalue process, we need some results on the perturbation theory of matrix (cf.[19]) and the delta method in large deviations ([9]).

Outline

Theorem 3.1

(1). Set $Y_t^\varepsilon := \frac{X_t^\varepsilon - X_t^0}{\sqrt{\varepsilon}}$*. Then* $Y^\varepsilon = (Y_t^\varepsilon)_{t\in[0,T]}$ *converges in probability on* $C_0([0,T],\mathcal{S}_m)$ *to the matrix process* Y^0*, determined by*

$$\begin{cases} dY_t^0 = \sqrt{X_t^0}dB_t + dB_t^T\sqrt{X_t^0}; \\ Y_0^0 = 0. \end{cases} \tag{3.4}$$

(2). $Z^\varepsilon = (Z_t^\varepsilon)_{t\in[0,T]}$ *obeys a large deviation principle on the space* $C_0([0,T],\mathcal{S}_m)$ *with speed* $h^2(\varepsilon)$ *and with good rate function*

$$I(\varphi) = \begin{cases} \frac{1}{2}\sum_{i,j=1}^m \int_0^T \frac{\dot{\varphi}_{ij}^2(t)}{x_{ii} + x_{jj} + 2\delta_{ij}x_{ii} + 2(1+\delta_{ij})\rho t}dt, & \varphi \in \mathcal{I}; \\ +\infty, & \text{otherwise}, \end{cases} \tag{3.5}$$

where $\mathcal{I} = \left\{\varphi \in C_0([0,T],\mathcal{S}_m) : \varphi_{ij} \in H, \forall i,j = 1,\cdots,m\right\}$ *and H is the Cameron-Martin space,.*

Outline

Let $\lambda_1^\varepsilon(t), \cdots, \lambda_m^\varepsilon(t)$, $t \in [0, T]$ be the ordered eigenvalues processes of X^ε. Then $\lambda_1^\varepsilon(t) < \cdots < \lambda_m^\varepsilon(t)$, $t \in [0, T]$, and it is the solution of the following SDE (cf. [4]):

$$\begin{cases} d\lambda_i^\varepsilon(t) = 2\sqrt{\varepsilon\lambda_i^\varepsilon(t)}d\beta_i(t) + \left(\rho + \varepsilon\sum_{k\neq i}\frac{\lambda_i^\varepsilon(t)+\lambda_k^\varepsilon(t)}{\lambda_i^\varepsilon(t)-\lambda_k^\varepsilon(t)}\right)dt, \quad i = 1, \cdots, m; \\ \lambda_i^\varepsilon(0) = \lambda_i, \end{cases}$$

Theorem 3.2

Assume that the initial value $\mathbf{x} \in \widetilde{\mathcal{S}}_m^+$ *has m distinct eigenvalues* $0 < \lambda_1 < \cdots < \lambda_m$. *Let* $b_k = (b_{k1}, \cdots, b_{km})$ *be the normal eigenvector corresponding to* λ_k, $k = 1, \cdots, m$, *and denote by* $\lambda_k^0(t) = \lambda_k + \rho t$, $k = 1, \cdots, m$, *which are the eigenvalue processes of* $X_t^0 = \mathbf{x} + \rho t I_m$. *Then* $\left\{\frac{\lambda^\varepsilon(t)-\lambda^0(t)}{\sqrt{\varepsilon}h(\varepsilon)} := \left(\frac{\lambda_1^\varepsilon(t)-\lambda_1^0(t)}{\sqrt{\varepsilon}h(\varepsilon)}, \cdots, \frac{\lambda_m^\varepsilon(t)-\lambda_m^0(t)}{\sqrt{\varepsilon}h(\varepsilon)}\right), t \in [0, T]\right\}$ *satisfies a large deviation principle on* $C_0([0, T], \mathbb{R}^m)$ *with speed* $h^2(\varepsilon)$ *and with good rate function*

$$I^\lambda(\phi) := \inf\left\{I(\psi); (b_1\psi b_1^T, \cdots, b_m\psi b_m^T) = \phi\right\}, \quad \phi \in C_0([0, T], \mathbb{R}^m).$$

Outline

1. Wigner matrix and Wishart process
   - Wigner matrix
   - Wishart process
2. Cramér type moderate deviations for eigenvalues of Wigner matrices
   - Cramér type moderate deviations
   - Sketch of the proof
   - Berry-Esseen bounds
   - Cramér type moderate deviations for covariance matrices
3. Moderate deviations for small perturbation Wishart processes
   - Small perturbation Wishart processes
   - Moderate deviations for small perturbation Wishart processes
   - Moderate deviations for the eigenvalue processes
   - Sketch of the proof

Key techniques:

- The delta method in large deviation theory ([9]).
- The perturbation theory of matrix (cf.[19]).

Lemma 3.1 (The delta method, Theorem 3.1 in [9])

Let $\mathcal{X}$ and $\mathcal{Y}$ be two metrizable linear topological spaces. Let $\Phi\colon \mathcal{D}_\Phi \subset \mathcal{X} \to \mathcal{Y}$ be Hadamard-differentiable at θ tangentially to $\mathcal{D}_0$, where $\mathcal{D}_\Phi$ and $\mathcal{D}_0$ are two subsets of $\mathcal{X}$. Let $X_n\colon \Omega_n \to \mathcal{D}_\Phi$, $n \geq 1$, be a sequence of maps and let r_n, $n \geq 1$, be a sequence of positive real numbers satisfying $r_n \to \infty$.
If $\left\{r_n(X_n - \theta), n \geq 1\right\}$ satisfies a large deviation principle with speed $\lambda(n)$ and with good rate function I and $\{I < \infty\} \subset \mathcal{D}_0$, then $\left\{r_n(\Phi(X_n) - \Phi(\theta)), n \geq 1\right\}$ satisfies a large deviation principle with speed $\lambda(n)$ and with good rate function $I_{\Phi'_\theta}$, where

$$I_{\Phi'_\theta}(y) = \inf\left\{I(x) : \Phi'_\theta(x) = y\right\}, \quad y \in \mathcal{Y}.$$

Lemma 3.2 (Gerschgorin's Theorem, cf. [19], Chapter IV, Theorem 2.1)

For $A = (a_{ij}) \in \mathbb{C}^{m\times m}$, *let* $\alpha_i = \sum_{j\neq i} |a_{ij}|$. *Then,*

$$\mathcal{L}(A) \subset \bigcup_{i=1}^{m} \mathcal{G}_i(A) := \bigcup_{i=1}^{m} \left\{ z \in \mathbb{C} : \left| z - a_{ii} \right| \leq \alpha_i \right\}.$$

Moreover, if l *of the Gerschgorin disks* $\mathcal{G}_i(A), i = 1, \cdots, m$, *are isolated from the other* $m - l$ *disks, then there are precisely* l *eigenvalues of* A *in their union.*

Let Ψ denote the map from $\mathcal{S}_m$ to $\mathbb{R}^m$ such that

$$\Psi(A) = (\mu_1, \cdots, \mu_m),$$

where $\mu_1 < \cdots < \mu_m$ denote the eigenvalues of A. Define $\Phi : C([0,T], \mathcal{S}_m) \to C([0,T], \mathbb{R}^m)$ as follows:

$$\Phi(Z)(t) = \Psi(Z(t)), \quad t \in [0,T]. \tag{3.6}$$

Lemma 3.3

Let $\psi \in C([0,T], \mathcal{S}_m)$ be given. Set $\|\psi\| = \sup_{t\in[0,T]} \|\psi_t\|_{HS}$ and $\delta = 1 \wedge \min_{2\le k\le m}(\lambda_k - \lambda_{k-1})$. Then for all sequences ε_n converging to $0+$ and satisfying $0 < \varepsilon_n \le \frac{\delta}{16m(\|\psi\|+1)}$, and $\psi^{(n)} \in C([0,T], \mathcal{S}_m)$ converging to ψ in $C([0,T], \mathcal{S}_m)$ and satisfying $\|\psi^{(n)}\| \le 2\|\psi\|$,

$$\max_{1\le k\le m} \sup_{t\in[0,T]} \left| \Phi\left(X^0 + \varepsilon_n \psi^{(n)}\right)_k (t) - \lambda_k^0(t) - \varepsilon_n b_k \psi_t^{(n)} b_k^T \right| \le \frac{32 m \varepsilon_n^2 \|\psi\|^2}{\delta}. \tag{3.7}$$

Thank You!

G.W. Anderson, A. Guionnet, and O. Zeitouni.
An introduction to random matrices, volume 200.
Cambridge University Press, 2010.

G.B. Arous and A. Guionnet.
Large deviations for Wigner's law and Voiculescu's non-commutative entropy.
Probability Theory and Related Fields, 108(4):517–542, 1997.

Z.D. Bai and J.W. Silverstein.
Spectral analysis of large dimensional random matrices.
Springer-Verlag, 2010.

M.F. Bru.
Wishart processes.
Journal of Theoretical Probability, 4(4):725–751, 1991.

L.H.Y. Chen, L. Goldstein, Q.M. Shao
Normal approximation by Stein's method
Springer-Verlag, 2011.

A. Dembo, A. Guionnet, and O. Zeitouni.
Moderate deviations for the spectral measure of certain random matrices.
In *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, volume 39, pages 1013–1042. Elsevier, 2003.

H. Döring and P. Eichelsbacher.
Moderate deviations via cumulants.
Arxiv preprint arXiv:1012.5027, 2010.

C. Donati-Martin.
Large deviations for wishart processes.
Probability and Mathematical Statistics, 28(2):325–343, 2008.

F.Q. Gao and X.Q. Zhao.
Delta method in large deviations and moderate deviations for estimators.
The Annals of Statistics, 39(2):1211–1240,2011.

A. Guionnet,

Large Random Matrices: Lectures on Mcroscopic Asymptotics. LNM 1957, Springer, 2009.

J. Gustavsson.
Gaussian fluctuations of eigenvalues in the GUE.
In *Annales de l'Institut Henri Poincare. Probability and Statistics*, volume 41, pages 151–178. Elsevier, 2005.

F. Hiai and D. Petz.
Eigenvalue density of the Wishart matrix and large deviations.
Infinite Dimensional Analysis, Quantum Probability and Related Topics, 1(4):633–646, 1998.

Y.T. Ma, R. Wang, and L.M. Wu,
Moderate deviation principle for dynamical systems with small random perturbation.
arXiv:1107.3432, July 2011.

M. Maïda.

Large deviations for the largest eigenvalue of rank one deformations of Gaussian ensembles.
Electronic Journal of Probability, 12:1131–1150, 2007.

M.L. Mehta.
Random matrices, volume 142.
Academic Press, 2004.

S. O'Rourke.
Gaussian fluctuations of eigenvalues in Wigner random matrices.
Journal of Statistical Physics, 138(6):1045–1066, 2010.

V.V. Petrov.
Limit theorems of probability theory: Sequences of independent random variables.
Clarendon Press (Oxford and New York), 1995.

L. Saulis and V.A. Statulevicius.
Limit theorems for large deviations, volume 73.
Springer, 1991.

- G.W. Stewart and J. Sun,
  Matrix Perturbation Theory.
  Academic Press, 1990.

- Z. Su.
  Gaussian fluctuations in complex sample covariance matrices.
  Electronic Journal of Probability, 11:1284–1320, 2006.

- T. Tao and V. Vu.
  Random matrices: Universality of local eigenvalue statistics up to the edge.
  Communications in Mathematical Physics, 298(2):549–572, 2010.

- T. Tao and V. Vu.
  Random matrices: Universality of local eigenvalue statistics.
  Acta Mathematica, pages 1–78, 2011.

- P. Vivo, S.N. Majumdar, and O. Bohigas.
  Large deviations of the maximum eigenvalue in Wishart random matrices.

Journal of Physics A: Mathematical and Theoretical, 40:4317–4337, 2007.

E.P. Wigner.
Characteristic vectors of bordered matrices with infinite dimensions.
Annals of Mathematics, 62(3):548–564, 1955.

E.P. Wigner.
On the distribution of the roots of certain symmetric matrices.
Annals of Mathematics, 67(2):325–327, 1958.