Lecture 31

Lecture date: Nov. 7, 2007 *Scribe: Anand Sarwate*

1 Gaussian concentration recap

If (W, T) is a pair of random variables such that

$$\mathbf{E}[W\varphi(W)] = \mathbf{E}[T\varphi'(W)] \tag{1}$$

for all Lipschitz φ, then for any $\sigma^2 > 0$

$$d_{TV}\left(\mathcal{L}(W), \mathcal{N}(0, \sigma^2)\right) \leq \frac{2\mathbf{E}|T - \sigma^2|}{\sigma^2}\ , \tag{2}$$

where $\mathcal{L}(W)$ is the law of W. In particular, if $\sigma^2 = \mathrm{Var}(W)$, it is easy to see that $\mathbf{E}[T] = \sigma^2$, and hence

$$d_{TV}\left(\mathcal{L}(W), \mathcal{N}(0, \sigma^2)\right) \leq \frac{2\sqrt{\mathrm{Var}(T)}}{\sigma^2}\ . \tag{3}$$

So how do we plan to use this? We have the following canonical construction: let $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ be a vector of iid $\mathcal{N}(0, 1)$ random variables and $f : \mathbb{R}^n \to \mathbb{R}$ be an absolutely continuous function. Then for $W = f(\mathbf{X})$ with $\mathbf{E}[W] = 0$, $\mathbf{E}[W^2] < \infty$ we have

$$\mathbf{E}[W\varphi(W)] = \mathbf{E}[T\varphi'(W)]\ , \tag{4}$$

for all Lipschitz φ, where

$$T = \int_0^1 \frac{1}{2\sqrt{t}} \sum_{i=1}^{n} \frac{\partial f}{\partial X_i}(\mathbf{X}) \frac{\partial f}{\partial X_i}(\mathbf{X}_t) dt \tag{5}$$

where $\mathbf{X}_t = \sqrt{t}\ \mathbf{X} + \sqrt{1-t}\ \mathbf{Y}$ and $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$ are also iid $\mathcal{N}(0, 1)$.

With this tool we proved the Gaussian Poincaré inequality and the Gaussian concentration inequality. Today we will start a method for obtaining normal approximations for quite complicated functions. For example, we will look at linear statistics of the eigenvalues of random matrices.

2 A CLT for functions of Gaussians

To get a CLT we first need to prove the concentration of T given by (5). Clearly, we can replace T by the conditional expectation $T(\mathbf{x}) = \mathbf{E}[T|\mathbf{X} = \mathbf{x}]$. This requires some ugly but straightforward calculation[1]. We begin by writing $T(\mathbf{x})$:

$$T(x_1, x_2, \ldots, x_n) = \int_0^1 \frac{1}{2\sqrt{t}} \mathbf{E}\left[\sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{x}) \frac{\partial f}{\partial x_i}(\sqrt{t}\ \mathbf{x} + \sqrt{1-t}\ \mathbf{Y})\right] dt \tag{6}$$

Letting $\sigma^2 = \mathrm{Var}(W)$, we have

$$d_{TV}\left(\mathcal{L}(W), \mathcal{N}(0,\sigma^2)\right) \le \frac{2}{\sigma^2}\sqrt{\mathrm{Var}(T(\mathbf{x}))} \tag{7}$$

By the Poincaré inequality,

$$\mathrm{Var}(T(\mathbf{X})) \le \mathbf{E}\left\|\nabla T(\mathbf{X})\right\|^2 \tag{8}$$

These give what one might call the "2nd order Poincaré inequalities."

Continuing with the computation:

$$\frac{\partial T}{\partial x_i}(\mathbf{x}) = \int_0^1 \frac{1}{2\sqrt{t}} \mathbf{E}\Bigg[\underbrace{\sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) \frac{\partial f}{\partial x_j}(\sqrt{t}\ \mathbf{x} + \sqrt{1-t}\ \mathbf{Y})}_{A_i(t)} \tag{9}$$

$$+\sqrt{t}\underbrace{\sum_{j=1}^n \frac{\partial f}{\partial x_j}(\mathbf{x}) \frac{\partial^2 f}{\partial x_i \partial x_j}(\sqrt{t}\ \mathbf{x} + \sqrt{1-t}\ \mathbf{Y})}_{B_i(t)}\Bigg] dt\ . \tag{10}$$

What we really want to bound is the sum of squares of this expression. Using the inequality $(a+b)^2 \le 2a^2 + 2b^2$ and Jensen's inequality, we get

$$\sum_{i=1}^n \left(\frac{\partial T}{\partial x_i}(\mathbf{x})\right)^2 \le 2\sum_{i=1}^n \left(\int_0^1 \frac{1}{2\sqrt{t}} \mathbf{E}[A_i(t)]dt\right)^2 + 2\sum_{i=1}^n \left(\int_0^1 \frac{1}{2}\mathbf{E}[B_i(t)]dt\right)^2 \tag{11}$$

$$\le 2\mathbf{E}\int_0^1 \frac{1}{2\sqrt{t}} \sum_{i=1}^n A_i(t)^2 dt + 2\mathbf{E}\int_0^1 \frac{1}{4}\sum_{i=1}^n B_i(t)^2 dt\ . \tag{12}$$

[1] You should be used to this by now!

Turning to the first term,

$$\sum_{i=1}^{n} A_i(t)^2 = \sum_{i=1}^{n}\left(\sum_{j=1}^{n}\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x})\frac{\partial f}{\partial x_j}(\sqrt{t}\ \mathbf{x}+\sqrt{1-t}\ \mathbf{Y})\right)^2 \tag{13}$$

$$= \left\|\text{Hess}\, f(\mathbf{x})\cdot\nabla f(\sqrt{t}\ \mathbf{x}+\sqrt{1-t}\ \mathbf{Y})\right\|^2 \tag{14}$$

$$\leq \|\text{Hess}\, f(\mathbf{x})\|^2\cdot\left\|\nabla f(\sqrt{t}\ \mathbf{x}+\sqrt{1-t}\ \mathbf{Y})\right\|^2 . \tag{15}$$

We can get a similar bound for $B_i(t)$.

Note that in computing $\mathbf{E}\,\|\nabla T(\mathbf{X})\|$ we will encounter terms that can be bounded using the Cauchy-Schwarz inequality and the fact that $\mathbf{X} \stackrel{d}{=} \mathbf{X}_t$.

$$\mathbf{E}\left[\|\text{Hess}\, f(\mathbf{X})\|^2\cdot\|\nabla f(\mathbf{X}_t)\|^2\right] \leq \left(\mathbf{E}\,\|\text{Hess}\, f(\mathbf{X})\|^4\right)^{1/2}\left(\mathbf{E}\,\|\nabla f(\mathbf{X}_t)\|^4\right)^{1/2} . \tag{16}$$

Thus

$$\text{Var}(T(\mathbf{X})) \leq \mathbf{E}\,\|\nabla T(\mathbf{X})\| \tag{17}$$

$$\leq \left(2\int_0^1 \frac{1}{2\sqrt{t}}dt + 2\int_0^1 \frac{1}{4}dt\right)\left(\mathbf{E}\,\|\text{Hess}\, f(\mathbf{X})\|^4\,\mathbf{E}\,\|\nabla f(\mathbf{X}_t)\|^4\right)^{1/2} \tag{18}$$

$$= \frac{5}{2}\left(\mathbf{E}\,\|\text{Hess}\, f(\mathbf{X})\|^4\,\mathbf{E}\,\|\nabla f(\mathbf{X}_t)\|^4\right)^{1/2} . \tag{19}$$

We then have

$$d_{TV}\left(\mathcal{L},\mathcal{N}(0,\sigma^2)\right) \leq \frac{2}{\sigma^2}\sqrt{\frac{5}{2}}\left(\mathbf{E}\,\|\text{Hess}\, f(\mathbf{X})\|^4\,\mathbf{E}\,\|\nabla f(\mathbf{X}_t)\|^4\right)^{1/4} \tag{20}$$

$$= \frac{\sqrt{10}}{\sigma^2}\left(\mathbf{E}\,\|\text{Hess}\, f(\mathbf{X})\|^4\,\mathbf{E}\,\|\nabla f(\mathbf{X}_t)\|^4\right)^{1/4} . \tag{21}$$

We have proved the following

Theorem 1 *If $W = f(X_1, X_2, \ldots, X_n)$ where $X = (X_1, X_2, \ldots, X_n)$ is a vector of iid $\mathcal{N}(0,1)$ random variables with $\mathbf{E}[W] = 0$, $\mathbf{E}[W^2] = \sigma^2$, and $f \in C^2(\mathbb{R}^n)$, then*

$$d_{TV}\left(\mathcal{L}(W),\mathcal{N}(0,\sigma^2)\right) \leq \frac{\sqrt{10}}{\sigma^2}\left(\mathbf{E}\,\|\text{Hess}\, f(\mathbf{X})\|^4\,\mathbf{E}\,\|\nabla f(\mathbf{X}_t)\|^4\right)^{1/4} . \tag{22}$$

Exercise 2 *Improve this theorem so that it doesn't have any 4th powers.*

3 Looking forward : eigenvalues of random matrices

What sort of problems can we tackle with this machinery? Suppose

$$(X_{ij})_{1 \le i,j < \infty} \tag{23}$$

are iid $\mathcal{N}(0,1)$ random variables and let

$$A_n = \frac{1}{\sqrt{n}} (X_{ij})_{1 \le i,j < \infty} \ . \tag{24}$$

This is sometimes called the real Ginibré ensemble. The eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of A_n are approximately uniformly distributed on the unit disc, in the following sense:

$$\frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i} \xrightarrow{a.s.} \text{Uniform(unit disc in } \mathbb{C}) \ . \tag{25}$$

We can look at sums of the form

$$\sum_{i=1}^{n} f(\lambda_i) \ , \tag{26}$$

for some function $f : \mathbb{C} \to \mathbb{C}$. It turns out that under very general conditions on f, this is asymptotically Gaussian, meaning

$$\sum_{i=1}^{n} f(\lambda_i) - \mathbf{E}\left[\sum_{i=1}^{n} f(\lambda_i) \right] \tag{27}$$

converges in law. For symmetric random matrices, Sinai and Soshnikov proved this in 1998.

We will conclude with a brief chronology of the relevant results.

1. **Jonsson, D. Some limit theorems for the eigenvalues of a sample covariance matrix J. Multivariate Anal. 12 1–38 (1982).**
   Discusses sample covariance or Wishart matrices, which are of the form $A^T A$, where A is a matrix whose rows are sample data points.

2. **Ya. Sinaĭ, A. Soshnikov, Central limit theorem for traces of large random symmetric matrices, Bol. Soc. Brasil. Mat., 29, No. 1, 1-24 (1998).**

3. **Ya. Sinaĭ, A. Soshnikov, A refinement of Wigner's semicircle law in a neighborhood of the spectrum edge for random symmetric matrices, Functional Anal. Appl. 32, No. 2, (1998).**
   These papers prove a refinement and CLT for Wigner matrices, which are symmetric random matrices.

4. **Johansson, K. On fluctuations of eigenvalues of random Hermitian matrices. Duke Math. J. 91 151–204 (1998).**
This paper studies matrices whose entries have joint density proportional to $\exp(-nTr(V(A)))$, where V is a polynomial.

5. **Diaconis, P. and Evans, S.N. Linear functionals of eigenvalues of random matrices. Trans. Amer. Math. Soc. 353 2615–2633 (2001).**
This paper studies random unitary matrices and uses connections to symmetric functions.

6. **Chatterjee, S. Fluctuations of eigenvalues and second order Poincaré inequalities. arXiv:0705.1224v2 [math.PR].**
This will be our plan for the next few lectures.