36-788 – Topics in High Dimensional Statistics I **Fall 2015**

Lecture 12: Oct 08

Lecturer: Alessandro Rinaldo *Scribes: Jisu KIM*

Note: *LaTeX template courtesy of UC Berkeley EECS dept.*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. They may be distributed outside this class only with the permission of the Instructor.*

12.1 Modified Logarithmic Sobolev Inequality

Proposition 12.1 [1] *Let $\phi(x) = e^x - x - 1$ for $x \in \mathbb{R}$ and $\lambda \in \mathbb{R}$, and let Z be an integrable random variable with $Z = f(\overbrace{X_1, \cdots, X_n}^{indep})$. Then,*

$$\begin{aligned} Ent(e^{\lambda Z}) &= \lambda \mathbb{E}\left[Z e^{\lambda Z}\right] - \mathbb{E}\left[e^{\lambda Z}\right] \log \mathbb{E}\left[e^{\lambda Z}\right] \\ &\leq \sum_{i=1}^{n} \mathbb{E}\left[e^{\lambda Z} \phi(-\lambda(Z - Z_i))\right] \end{aligned}$$

where each Z_i is an (arbitrary) function of $X^{(i)} = \{X_j,\ j \neq i\}$.

Proof: Based on sub-additivity of Ent, recall ($Y \geq 0$ random variable)

$$Ent(Y) \leq \mathbb{E}\left[Y \log Y - Y \log \mu - (Y - \mu)\right]$$

[2] for any $\mu > 0$. Use this conditionally on $X^{(i)}$,

$$\begin{aligned} Ent^{(i)}\left[e^{\lambda Z}\right] &\leq \mathbb{E}^{(i)}\left[e^{\lambda Z}\left(\log e^{\lambda Z} - \log e^{\lambda Z_i}\right) - (e^{\lambda Z} - e^{\lambda Z_i})\right] \\ &= \mathbb{E}^{(i)}\left[e^{\lambda Z} \phi(-\lambda(Z - Z_i))\right] \end{aligned}$$

where $\mathbb{E}^{(i)}$ is a conditional expectation given $X^{(i)}$. This is from plugging $Y \to e^{\lambda Z}$ and $\mu \to e^{\lambda Z_i}$ which is constant given $X^{(i)}$. Now use

$$Ent(e^{\lambda Z}) \leq \sum_{i=1}^{n} \mathbb{E}\left(Ent^{(i)}\left[e^{\lambda Z}\right]\right)$$

to get result. ∎

Application 1.

Let $Z = f(\underbrace{X_1, \cdots, X_n}_{\text{indep}})$ be such that $\sum_{i=1}^{n}(Z - Z_i)^2 \leq v^2$ a.s., where $Z_i = \inf_{x_i} f\left(X_1, \cdots, X_{i-1}, x_i, X_{i+1}, \cdots, X_n\right)$ (Z_i is measurable function of $X^{(i)}$). Then $\forall t > 0$,

$$P\left(Z - \mathbb{E}[Z] \geq t\right) \leq \exp\left\{-\frac{t^2}{2v^2}\right\},$$

[1] Proposition 4 in [M2006], p. 6, Section 3.4 in [M2012], p. 9
[2] See formulation of entropy in Lecture 10(Oct 01)

so Z has Gaussian tail with v^2 as variance factor.

Proof: $\phi(x) = e^x - x - 1$, and check $\phi(-x) \leq \frac{x^2}{2}$ $\forall x$. Then $\forall \lambda > 0$,

$$\begin{aligned} Ent(e^{\lambda Z}) &\leq \mathbb{E}\left[e^{\lambda Z}\sum_{i=1}^{n}\frac{\lambda^2}{2}(Z - Z_i)^2\right] \\ &\leq \mathbb{E}\left[e^{\lambda Z}\right]\frac{\lambda^2}{2}v^2 \end{aligned}$$

Now use Herbst argument![3] ■

Remark.

- By taking $Z_i = \sup_{x_i'} f(X_1, \cdots, x_i', \cdots, X_n)$, with condition $\sum_{i=1}^{n}(Z - Z_i)^2 \leq v^2$ a.s., then $\forall t > 0$, $P\left(Z - \mathbb{E}[Z] \leq -t\right) \leq \exp\left\{-\frac{t^2}{2v^2}\right\}$, so

$$P\left(|Z - \mathbb{E}[Z]| \geq t\right) \leq 2e^{-\frac{t^2}{2v^2}}.$$

- For bounded difference inequality[4],

$$\sum_{i=1}^{n} \sup_{x_1, \cdots, x_n, x_1', \cdots, x_n'} \left(f(x_1, \cdots, x_n) - f(x_1, \cdots, x_i', \cdots, x_n)\right)^2 \leq \frac{v^2}{4},$$

  where sup is inside. For this inequality, we require

$$\sup_{x_1, \cdots, x_n, x_1', \cdots, x_n'} \sum_{i=1}^{n} \left(f(x_1, \cdots, x_n) - f(x_1, \cdots, x_i', \cdots, x_n)\right)^2 \leq v^2,$$

  where sup is outside, so this condition is weaker.

- Let $\underset{n\times n \text{ symmetric}}{A} = (X_{i,j})_{1\leq i\leq j\leq n}$ be independent, $|X_{ij}| < 1$ a.s. Let $Z = \lambda_1$, i.e. largest eigenvalue of A. Then

$$\sum_{i,j}(Z - Z_{i,j})^2 \leq 16,$$

  where $Z_{i,j}$ is infimum of Z over i, j coordinate of A. Hence

$$P\left(\lambda_1 - \mathbb{E}[\lambda_1] \geq t\right) \leq \exp\left\{-\frac{t^2}{32}\right\}.$$

  This will be hard to obtain using McDiarmid[5].

Application 2.

Let $X_1, \cdots, X_n$ independent, taking values in $[0, 1]$. Let $f : [0, 1]^n \to \mathbb{R}$ Lipschitz and separately convex. Set $Z = f(X_1, \cdots, X_n)$. Then

$$P\left(Z - \mathbb{E}[Z] \geq t\right) \leq \exp\left\{-\frac{t^2}{2}\right\}.$$

---

[3]See Lecture 10(Oct 01)
[4]See Lecture 10(Oct 01)
[5]See Lecture 10(Oct 01)

Proof: Assume partial derivatives exist. Let $Z_i = \inf_{x_i'} f(X_1, \cdots, x_i', \cdots, X_n)$ function of $X^{(i)}$. Let $\tilde{X}_i$ (function of $X^{(i)}$) be the value of x_i' at which infimum is achieved. Then

$$\begin{aligned}\sum_{i=1}^{n}(Z - Z_i)^2 &\leq \sum_{i=1}^{n}\left[\frac{\partial f(X)}{\partial x_i}(X_i - \tilde{X}_i)\right]^2 \text{ (convexity)}\\ &\leq \sum_{i=1}^{n}\left[\frac{\partial f}{\partial x_i}(X)\right]^2 = \|\nabla f(X)\|^2\\ &\leq 1 \text{ (by Lipschitz assumption)}\end{aligned}$$

Use theorem with $v^2 = 1$. ■

12.2 Self bounding functions

In following, $Z = f(\underbrace{X_1, \cdots, X_n}_{\text{indep}})$, and $h(\mu) = (1+\mu)\log(1+\mu) - \mu$ for $\mu \geq -1$, and $\phi(v) = e^v - v - 1 = \sup_{\mu \geq -1}\{\mu v - h(\mu)\}$.

Theorem 12.2 [6] *If f has SBP*[7], $\forall \lambda \in \mathbb{R}$,

$$\log E\left[e^{\lambda(Z - \mathbb{E}[Z\})}\right] \leq \phi(\lambda)\mathbb{E}[Z]$$

by Chernoff calculation, this implies then $\forall t \geq 0$,

$$P(Z - \mathbb{E}[Z] \geq t) \leq \exp\left\{-h\left(\frac{t}{\mathbb{E}[Z]}\right)\mathbb{E}[Z]\right\}$$

and $\forall t \in (0, \mathbb{E}[Z]]$,

$$P(Z \leq \mathbb{E}[Z] - t) \leq \exp\left\{-h\left(-\frac{t}{\mathbb{E}[Z]}\right)\mathbb{E}[Z]\right\}$$

which is same Chernoff bound as if $Z \sim$Poisson. Since

$$h(\mu) \geq \frac{\mu^2}{2 + \frac{2}{3}\mu}, \ \mu \geq 0$$

we get

$$\begin{aligned}P(Z - \mathbb{E}[Z] \geq t) &\leq \exp\left\{-\frac{t^2}{2\mathbb{E}[Z] + \frac{2}{3}t}\right\} && t \geq 0\\ P(Z \leq \mathbb{E}[Z] - t) &\leq \exp\left\{-\frac{t^2}{2\mathbb{E}[Z]}\right\} && 0 < t < \mathbb{E}[Z]\end{aligned}$$

Remark. Think of chi square. Upper tail is thicker.

Proof: $\phi(x) = e^x - x - 1$, then notice that for $\mu \in (0,1)$ and $\lambda \in \mathbb{R}$,

$$\phi(-\lambda\mu) \leq \mu\phi(-\lambda).$$

[6]Theorem 19 in [M2012], p. 18
[7]See Lecture 7(Sep 22) and Lecture 9(Sep 29)

So, since $Z - Z_i \in [0,1]$,

$$\phi(-\lambda(Z - Z_i)) \le (Z - Z_i)\phi(-\lambda).$$

By modified log sobolev inequality,

$$\begin{aligned} Ent(e^{\lambda Z}) &\le \mathbb{E}\left[e^{\lambda Z}\sum_{i=1}^{n}\phi\left(-\lambda(Z - Z_i)\right)\right] \\ &\le \mathbb{E}\left[e^{\lambda Z}\phi(-\lambda)\underbrace{\sum_{i=1}^{n}(Z - Z_i)}_{\le Z}\right] \\ &\le \phi(-\lambda)\mathbb{E}\left[e^{\lambda Z}Z\right] \end{aligned}$$

Rest of the proof is boring calculus. ■

Example.

1. Let $\mathbb{G}(n,p)$ be an Erdős-Rényi random graph, i.e. edge between node i and j is $X_{ij} \sim Bernoulli(p)$ independently. Let d_i degree on node i. Then $d_i \sim Bin(n-1,p)$, and $D = \max_i d_i = f(X_{i,j}, i < j)$ satisfies SBP.
2. Conditional Rademacher averages. Let $Z = \mathbb{E}\left[\sup_{t\in T}\sum_{i=1}^{n}\epsilon_i x_{i,t}\right]$, where $x_i \in [0,1]^T$, $x_{i,t}$ t-component of x_i, and $\epsilon_1, \cdots, \epsilon_n \overset{iid}{\sim}$Rademacher. Then Z has SBP. (Homework)

12.3 Exponential Efron Stein inequality

Efron-Stein inequality states $V[Z] \le V[V_+] = V[V_-]$, where $V_+ = \sum_{i=1}^{n}\mathbb{E}\left[(Z - Z_i')_+^2\right]$ and $V_- = \sum_{i=1}^{n}\mathbb{E}\left[(Z - Z_i')_-^2\right]$.

Theorem 12.3 [8] *Let $\theta, \lambda > 0$ such that $\lambda\theta > 1$. Assume $\mathbb{E}\left[e^{\frac{\lambda V_\pm}{\theta}}\right] < \infty$. Then*

$$\log \mathbb{E}\left[e^{\lambda(Z - \mathbb{E}[Z])}\right] \le \frac{\lambda\theta}{1 - \lambda\theta}\mathbb{E}\left[e^{\frac{\lambda V_+}{\theta}}\right]$$

and

$$\log \mathbb{E}\left[e^{\lambda(Z - \mathbb{E}[Z])}\right] \le \frac{\lambda\theta}{1 - \lambda\theta}\mathbb{E}\left[e^{\frac{\lambda V_-}{\theta}}\right].$$

Applications

1. [9] If $V_+, V_- \le c$ a.e. for some $c > 0$, then

$$P\left(|Z - \mathbb{E}[Z]| \ge t\right) \le \exp\left\{-\frac{t^2}{4c}\right\}.$$

[8] Theorem 2 in [BLM2003], p. 1585
[9] Corollary 3 in [BLM2003], p. 1585-1586

2. [10] If $V_+ \leq aZ + b$, then nice exponential inequality.

3. Weakly (a, b) SB(self bounding) functions(weakening of SBP):

$$\sum_{i=1}^{n}(Z - Z_i)^2 \leq aZ + b,\ a, b \geq 0$$

then still nice exponential inequality.

Reference

[BLM2003] S. Boucheron and G. Lugosi and P. Massart: Concentration inequalities using the entropy method (2003), Annals of Probability, 31, 1583-1614.

[M2006] Maurer, A. (2006). Concentration inequalities for functions of independent variables. Random Structures & Algorithms, 29, 121138.

[M2012] A. Maurer (2012). Thermodynamics and concentration, Bernoulli, 18(2), 434-454.

[10] Theorem 5 in [BLM2003], p. 1587