Lecture 14

Agenda: log-Sobolev and entropy method for general r.v.s

A A modified log-Sobolev inequality

We now derive a variant of LS ineq. which is valid for all distributions

Theorem. $X_1, \dots, X_n$ indep., $Z = f(X_1, \dots, X_n)$

Let $\phi(x) = e^x - x - 1$. Then, $\forall \lambda > 0$,

$$\text{Ent}(e^{\lambda Z}) \le \sum_{i=1}^n \mathbb{E}\left[e^{\lambda Z} \phi(-\lambda(Z - Z_i))\right],$$

where $Z_i = f_i(X^{(i)}) = f_i(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)$

and f_i is any function of $X^{(i)}$.

Proof. By the sub-additivity of entropy, it suffices to show

$$\text{Ent}_{(i)}[e^{\lambda Z}] = \mathbb{E}_{(i)}\left[\lambda Z e^{\lambda Z}\right] - \mathbb{E}_{(i)}[e^{\lambda Z}] \log \mathbb{E}_{(i)}[e^{\lambda Z}]$$
$$\le \mathbb{E}_{(i)}\left[e^{\lambda Z} \cdot \phi(-\lambda(Z - Z_i))\right]$$

Note that for $x \ge 0$, $u > 0$

$$x \log u + x - u \le x \log x.$$

Thus, $-\mathbb{E}_{(i)}[e^{\lambda Z}] \log \mathbb{E}_{(i)}[e^{\lambda Z}]$

$$\le -\mathbb{E}_{(i)}\left[e^{\lambda Z} \log e^{\lambda Z_i} + e^{\lambda Z} - e^{\lambda Z_i}\right]$$
$$= -\mathbb{E}_{(i)}\left[e^{\lambda Z}\left(\lambda Z_i + 1 - e^{-\lambda(Z - Z_i)}\right)\right]$$

$$\Rightarrow \mathrm{Ent}_{(i)}[e^{\lambda Z}] \le \mathbb{E}_{(i)}[e^{\lambda Z}(\lambda(Z-Z_i)+1-e^{-\lambda(Z-Z_i)})]$$
$$= \mathbb{E}_{(i)}[e^{\lambda Z}\phi(-\lambda(Z-Z_i))]. \qquad \blacksquare$$

The function $\phi(x) = e^x - x - 1$ is the log-moment generating function of a centralized Poi(1) r.v.

Properties of $\phi(x)$:

(i) $\boxed{\phi(-x) \le x^2/2, \quad \forall x > 0}$

Proof. $g(x) = \phi(-x)$. $g(0) = 0$, $g'(x) = -e^{-x}+1 \le x$ $\forall x > 0$.

Thus, $\phi(-x) \le x^2/2$. $\blacksquare$

(ii) $\boxed{\phi(\lambda x) \le \phi(\lambda)x^2 \quad \forall \lambda > 0 \text{ and } 0 \le x \le 1.}$

Proof. Consider $g(y) = \phi(y)/y^2$. It suffices to show that $g(\lambda x) \le g(\lambda)$, which in turn will follow if we show that g is nondecreasing. Towards that, note

$$g'(y) = \frac{1}{3!} + \frac{2y}{4!} + \frac{3y^2}{5!} + \cdots \ge 0. \qquad \blacksquare$$

$\boxed{B}$ Upper tail bound

Theorem. $Z = f(X_1, \ldots, X_n)$; $Z_i = \inf_{x_i'} f(X^{i-1}, x_i', X_{i+1}^n)$

Assume that $\exists v > 0$

$$\sum_{i=1}^n (Z - Z_i)^2 \le v.$$

Then $\forall t > 0$, $\mathbb{P}(Z - \mathbb{E}Z > t) \le e^{-t^2/2v}$.

Proof. Using the modified LSI and $\phi(-x) \le \frac{x^2}{2}, x>0$, ③

$$\mathrm{Ent}(e^{\lambda Z}) \le \sum_{i=1}^{n} \mathbb{E}\left[e^{\lambda Z}\phi(-\lambda\underbrace{(Z-Z_i)}_{\ge 0})\right], \quad \lambda>0$$

$$\le \sum_{i=1}^{n} \mathbb{E}\left[e^{\lambda Z}\frac{\lambda^2(Z-Z_i)^2}{2}\right]$$

$$\le \frac{\lambda^2}{2}.v.\mathbb{E}[e^{\lambda Z}].$$

The claim follows using Herbst's argument. ∎

Left-tail?

Assume $\sup_{\substack{x_1,\dots,x_n\\ x_1',\dots,x_n'}} \sum_{i=1}^{n}\left(f(x)-f(\tilde{x}^{(i)})\right)^2 \le v.$

Then, $\mathbb{P}(|Z-\mathbb{E}Z|>t) \le 2.e^{-t^2/2v}.$

Q. Why is this bound better than McDiarmid?

→ Recall our earlier analysis for $Z=$ max eig of A s.t.

A is symmetric real random matrix with upper diagonal entries X_{ij}, $1\le i\le j\le n$, which are iid and $|X_{ij}|\le 1$, a.s.

We could show: $\sum_{1\le i\le j\le n}(Z-Z_{ij}')^2 \le 16.$

But the BDP doesn't hold with reasonable c.

$\boxed{C}$ Lower tail bound

Theorem. $Z=f(X)$, $Z_i = \sup_{x_i'} f(X^{i-1}, x_i', X_{i+1}^n)$

Assume $\sum_{i=1}^{n}(Z_i-Z)^2 \le v$; and $Z_i - Z \le 1$, $1\le i\le n$.

Then, $\mathbb{P}(Z-\mathbb{E}Z \geq t) \leq e^{-v\,h(t/v)}$,

where $h(x) = (1+x)\log(1+x) - x\log x > -1$.

Proof. $\mathrm{Ent}(e^{\lambda Z}) \leq \sum_{i=1}^{n} \mathbb{E}\left[e^{\lambda Z}\phi(-\lambda(Z-Z_i))\right]$, $\lambda < 0$

[since $0 \leq Z_i - Z \leq 1$] $\quad \leq \sum_{i=1}^{n} \mathbb{E}\left[e^{\lambda Z}\phi(\lambda)(Z-Z_i)^2\right]$

$$= \phi(\lambda)\,\mathbb{E}\left[e^{\lambda Z}\sum_{i=1}^{n}(Z-Z_i)^2\right]$$

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

[Heuristically: like Herbst's argument, $\psi_{Z-\mathbb{E}Z}(\lambda) \leq \phi(\lambda)v$]

Let $g(\lambda) = \dfrac{\psi_Z(\lambda)}{\lambda}$.

$$g'(\lambda) = \frac{\mathbb{E}[Z e^{\lambda Z}]}{\lambda\,\mathbb{E}[e^{\lambda Z}]} - \frac{\log \mathbb{E}[e^{\lambda Z}]}{\lambda^2}$$

$$= \frac{1}{\lambda^2}\left(\frac{\mathbb{E}[\lambda Z e^{\lambda Z}] - \mathbb{E}[e^{\lambda Z}]\log \mathbb{E}[e^{\lambda Z}]}{\mathbb{E}[e^{\lambda Z}]}\right)$$

$$\leq \frac{\phi(\lambda)}{\lambda^2}v$$

Thus, $\dfrac{\psi_Z(\lambda)}{\lambda} - \psi_Z'(0) \leq \left(\displaystyle\int_0^{\lambda} \frac{\phi(x)}{x^2}\,dx\right).v$

Note that $\displaystyle\int_0^{\lambda}\frac{\phi(x)}{x^2}dx = \left(\frac{\lambda}{2!} + \frac{\lambda^2}{2.3!} + \cdots\right) \leq \frac{e^{\lambda}-1-\lambda}{\lambda}$, i.e.,

$$\frac{\psi_Z(\lambda)}{\lambda} - \mathbb{E}[Z] \leq \frac{\phi(\lambda)}{\lambda}.v$$

$$\Leftrightarrow \psi_{Z-\mathbb{E}}(\lambda) \leq \phi(\lambda).v. \qquad \blacksquare$$

Next: (I) Concentration of weakly self-bounding functions

(II) Relaxing $\sum_{i=1}^{n}(Z-Z_i)^2 \le v$, $Z_i - Z \le 1$.

(I). $\boxed{A}$ Upper tail bound for weakly self-bounding functions

Definition. A function $f: \mathcal{X}^n \to [0,\infty)$ is called weakly (a,b)-self bounding if there exist functions $f_i : \mathcal{X}^{(i)} \to [0,\infty)$ s.t. $\forall x \in \mathcal{X}^n$

$$\sum_{i=1}^{n} \left(f(x) - f_i(x^{(i)})\right)^2 \le a f(x) + b, \quad \forall x \in \mathcal{X}^n.$$

Theorem. If f is weakly (a,b)-self bounding s.t. $f(x) \ge f_i(x^{(i)})$ $\forall\, 1 \le i \le n$ and x, we have

$$\mathbb{P}\left(Z \ge \mathbb{E}Z + t\right) \le \exp\left(\frac{-t^2}{2(a\,\mathbb{E}Z + b + at/2)}\right).$$

Proof.

$$\lambda\, \mathbb{E}[Ze^{\lambda Z}] - \mathbb{E}[e^{\lambda Z}]\log \mathbb{E}[e^{\lambda Z}]$$

$$\le \sum_{i=1}^{n} \mathbb{E}\left[e^{\lambda Z}\, \phi(-\lambda(Z-Z_i))\right]$$

$$\le \frac{\lambda^2}{2} \sum_{i=1}^{n} \mathbb{E}\left[e^{\lambda Z}(Z-Z_i)^2\right]$$

$$\le \frac{\lambda^2}{2} \mathbb{E}\left[e^{\lambda Z}(aZ+b)\right]$$

Rearranging the terms, we get

$$\left(\frac{1}{\lambda} - \frac{a}{2}\right) \frac{\mathbb{E}[Z e^{\lambda Z}]}{\mathbb{E}[e^{\lambda Z}]} - \frac{\log \mathbb{E}[e^{\lambda Z}]}{\lambda^2} \le \frac{b}{2}$$

$$\Leftrightarrow \quad \frac{d}{d\lambda}\left(\left(\frac{1}{\lambda} - \frac{a}{2}\right) \log \mathbb{E}[e^{\lambda Z}]\right) \le \frac{b}{2}$$

$$\Rightarrow \left(\frac{1}{\lambda} - \frac{a}{2}\right) \log \mathbb{E}[e^{\lambda Z}] - \lim_{\lambda \to 0} \frac{\log \mathbb{E}[e^{\lambda Z}]}{\lambda}$$

$$+ \frac{a\lambda}{2}\mathbb{E}[Z] \le \frac{\lambda}{2}(b + a\mathbb{E}Z) =: \frac{\lambda v}{2}$$

$$\Leftrightarrow \left(\frac{1}{\lambda} - \frac{a}{2}\right) \log \mathbb{E}[e^{\lambda Z}] - \frac{\lambda \mathbb{E}[Z]}{\lambda}\left(1 - \frac{a\lambda}{2}\right) \le \frac{\lambda v}{2}$$

$$\Leftrightarrow \left(\frac{1}{\lambda} - \frac{a}{2}\right) \log \mathbb{E}[e^{\lambda(Z - \mathbb{E}Z)}] \le \frac{\lambda v}{2}$$

$$\Leftrightarrow \psi_{Z - \mathbb{E}Z}(\lambda) \le \frac{\lambda^2 v}{(1 - a\lambda/2)}. \qquad \blacksquare$$