Lecture 4: The Entropy Methods

Tianpei Xie

Jan. 19th., 2023

Contents

1 Logarithmic Sobolev Inequality

1.1 Logarithmic Sobolev Inequality for Bernoulli Distributions

- **Remark** (*Setting*)
  Consider a ***uniformly distributed binary vector*** $Z = (Z_1, \ldots, Z_n)$ on the hypercube $\{-1, +1\}^n$. In other words, the components of X are *independent*, *identically distributed* ***random sign (Rademacher) variables*** with $\mathbb{P}\{Z_i = -1\} = \mathbb{P}\{Z_i = +1\} = 1/2$ (i.e. *symmetric Bernoulli random variables*).

  Let $f : \{-1, +1\}^n \to \mathbb{R}$ be a real-valued function on ***binary hypercube***. $X := f(Z)$ is an induced real-valued random variable. Define $\widetilde{Z}^{(i)} = (Z_1, \ldots, Z_{i-1}, Z'_i, Z_{i+1}, \ldots, Z_n)$ be the sample Z with i-th component replaced by an *independent copy* Z'_i. Since $Z, \widetilde{Z}^{(i)} \in \{-1, +1\}^n$, $\widetilde{Z}^{(i)} = (Z_1, \ldots, Z_{i-1}, -Z_i, Z_{i+1}, \ldots, Z_n)$, i.e. *the i-th sign is **flipped***. Also denote the i-th *Jackknife sample* as $Z_{(i)} = (Z_1, \ldots, Z_{i-1}, Z_{i+1}, \ldots, Z_n)$ by *leaving out* the i-th component. $\mathbb{E}_{(-i)}[X] := \mathbb{E}\left[X | Z_{(i)}\right]$.

  Denote the i-th component of ***discrete gradient*** of f as

  $$\nabla_i f(z) := \frac{1}{2}\left(f(z) - f(\widetilde{z}^{(i)})\right)$$

  and $\nabla f(z) = (\nabla_1 f(z), \ldots, \nabla_n f(z))$

- **Remark** (***Jackknife Estimate of Variance***)
  Recall that *the **Jackknife estimate of variance***

  $$\begin{aligned}\mathcal{E}(f) &:= \mathbb{E}\left[\sum_{i=1}^{n}\left(f(Z) - \mathbb{E}_{(-i)}\left[f(\widetilde{Z}^{(i)})\right]\right)^2\right] \\ &= \frac{1}{2}\mathbb{E}\left[\sum_{i=1}^{n}\left(f(Z) - f(\widetilde{Z}^{(i)})\right)^2\right].\end{aligned}$$

  Using the notation of discrete gradient of f, we see that

  $$\mathcal{E}(f) := 2\mathbb{E}\left[\|\nabla f(Z)\|_2^2\right]$$

- **Remark** (***Entropy Functional***)
  Recall that the entropy functional for f is defined as

  $$H_\Phi(f(Z)) = \mathrm{Ent}(f) := \mathbb{E}\left[f(Z)\log f(Z)\right] - \mathbb{E}\left[f(Z)\right]\log\left(\mathbb{E}\left[f(Z)\right]\right).$$

- **Proposition 1.1** *(**Logarithmic Sobolev Inequality for Function of Rademacher Random Variables**). [Boucheron et al., 2013]*
  *If $f : \{-1, +1\}^n \to \mathbb{R}$ be an arbitrary real-valued function defined on the n-dimensional **binary hypercube** and assume that Z is **uniformly distributed** over $\{-1, +1\}^n$. Then*

  $$\begin{aligned}Ent(f^2) &\le \mathcal{E}(f) &(1)\\ \Leftrightarrow Ent(f^2(Z)) &\le 2\mathbb{E}\left[\|\nabla f(Z)\|_2^2\right] &(2)\end{aligned}$$

Proof: The key is to apply the tensorization property of Φ-entropy. Let $X = f(Z)$. By tensorization property,

$$\mathrm{Ent}(X^2) \le \sum_{i=1}^{n} \mathbb{E}\left[\mathrm{Ent}_{(-i)}(X^2)\right]$$

where $\mathrm{Ent}_{(-i)}(X^2) := \mathbb{E}_{(-i)}\left[X^2 \log X^2\right] - \mathbb{E}_{(-i)}\left[X^2\right] \log\left(\mathbb{E}_{(-i)}\left[X^2\right]\right)$.

It thus suffice to show that for all $i = 1, \ldots, n$,

$$\mathrm{Ent}_{(-i)}(X^2) \le \frac{1}{2}\mathbb{E}_{(-i)}\left[\left(f(Z) - f(\widetilde{Z}^{(i)})\right)^2\right].$$

Given any fixed realization of $Z_{(-i)}$, $X = f(Z) = \widetilde{f}(Z_i)$ can only takes two different values with equal probability. Call these two values a and b. See that

$$\mathrm{Ent}_{(-i)}(X^2) = \frac{1}{2}a^2 \log a^2 + \frac{1}{2}b^2 \log b^2 - \frac{1}{2}(a^2+b^2)\log\left(\frac{a^2+b^2}{2}\right)$$

$$\frac{1}{2}\mathbb{E}_{(-i)}\left[\left(f(Z) - f(\widetilde{Z}^{(i)})\right)^2\right] = \frac{1}{2}(a-b)^2.$$

Thus we need to show

$$\frac{1}{2}a^2 \log a^2 + \frac{1}{2}b^2 \log b^2 - \frac{1}{2}(a^2+b^2)\log\left(\frac{a^2+b^2}{2}\right) \le \frac{1}{2}(a-b)^2.$$

By symmetry, we may assume that $a \ge b$. Since $(|a| - |b|)^2 \le (a-b)^2$, without loss of generality, we may further assume that $a, b \ge 0$.

Define

$$h(a) := \frac{1}{2}a^2 \log a^2 + \frac{1}{2}b^2 \log b^2 - \frac{1}{2}(a^2+b^2)\log\left(\frac{a^2+b^2}{2}\right) - \frac{1}{2}(a-b)^2$$

for $a \in [b, \infty)$. $h(b) = 0$. It suffice to check that $h'(b) = 0$ and that h is concave on $[b, \infty)$. Note that

$$\begin{aligned} h'(a) &= a \log a^2 + 1 - a \log\left(\frac{a^2+b^2}{2}\right) - 1 - (a-b) \\ &= a \log \frac{2a^2}{(a^2+b^2)} - (a-b). \end{aligned}$$

So $h'(b) = 0$. Moreover,

$$h''(a) = \log \frac{2a^2}{(a^2+b^2)} + 1 - \frac{2a^2}{(a^2+b^2)} \le 0$$

due to inequality $\log(x) + 1 \le x$. ∎

- **Remark** (***Logarithmic Sobolev Inequality Stronger than Efron-Stein Inequality***). [Boucheron et al., 2013]
  Note that for f non-negative,

$$\mathrm{Var}(f(Z)) \le \mathrm{Ent}(f^2(Z)).$$

Thus *logarithmic Sobolev inequality* (1) implies

$$\mathrm{Var}(f(Z)) \leq \mathcal{E}(f)$$

which is *the Efron-Stein inequality.*

- **Corollary 1.2** ***(Logarithmic Sobolev Inequality for Function of Asymmetric Bernoulli Random Variables)****. [Boucheron et al., 2013]*
  If $f : \{-1, +1\}^n \to \mathbb{R}$ be an arbitrary real-valued function and $Z = (Z_1, \ldots, Z_n) \in \{-1, +1\}^n$ with $p = \mathbb{P}\{Z_i = +1\}$. Then

$$Ent(f^2) \leq \frac{1}{2} c(p) \mathcal{E}(f) \tag{3}$$

  where

$$c(p) = \frac{1}{1 - 2p} \log \frac{1 - p}{p}$$

  Note that $\lim_{p \to 1/2} c(p) = 2$.

1.2 Gaussian Logarithmic Sobolev Inequality

- **Proposition 1.3** ***(Gaussian Logarithmic Sobolev Inequality)****. [Boucheron et al., 2013]*
  Let $f : \mathbb{R}^n \to \mathbb{R}$ be a ***continuous differentiable*** *function and let $Z = (Z_1, \ldots, Z_n)$ be a vector of n* ***independent standard Gaussian*** *random variables. Then*

$$Ent(f^2(Z)) \leq 2\mathbb{E}\left[\|\nabla f(Z)\|_2^2\right]. \tag{4}$$

- **Remark** (***Gaussian Logarithmic Sobolev Inequality Stronger than Gaussian Poincaré Inequality***). [Boucheron et al., 2013]
  Recall that *the Gaussian Poincaré inequality*

$$\mathrm{Var}(f(Z)) \leq \mathbb{E}\left[\|\nabla f(Z)\|_2^2\right]$$

  We can show that for Gaussian random vectors Z,

$$2\mathrm{Var}(f(Z)) \leq \mathrm{Ent}(f^2(Z)).$$

  Thus *the Gaussian logarithmic Sobolev inequality* implies *the Gaussian Poincaré inequality.*

1.3 Logarithmic Sobolev Inequality for Binomial and Poisson Distributions

1.4 Logarithmic Sobolev Inequality for General Probability Measures

- **Definition** (***Logarithmic Sobolev Inequality for General Probability Measure***).
  A probability measure μ on $\mathbb{R}^n$ is said to satisfy the <u>***logarithmic Sobolev inequality***</u> for some constant $C > 0$ if

$$\mathrm{Ent}_\mu(f^2) \leq C\, \mathbb{E}_\mu\left[\|\nabla f\|_2^2\right] \tag{5}$$

holds for any ***continuous differentiable*** function $f: \mathbb{R}^n \to \mathbb{R}$. The left-hand side is called ***the entropy functional***, which is defined as

$$\begin{aligned}\mathrm{Ent}(f^2) &:= \mathbb{E}_\mu\left[f^2 \log f^2\right] - \mathbb{E}_\mu\left[f^2\right] \log \mathbb{E}_\mu\left[f^2\right] \\ &= \int f^2 \log\left(\frac{f^2}{\int f^2 d\mu}\right) d\mu.\end{aligned}$$

The right-hand side is defined as

$$\mathbb{E}_\mu\left[\|\nabla f\|_2^2\right] = \int \|\nabla f\|_2^2\, d\mu.$$

Thus we can rewrite *the logarithmic Sobolev inequality* in *functional form*

$$\int f^2 \log\left(\frac{f^2}{\int f^2 d\mu}\right) d\mu \le C \int \|\nabla f\|_2^2\, d\mu \tag{6}$$

- **Remark** (***Logarithmic Sobolev Inequality***)
  We can replace $f \to \sqrt{f}$, so that *the logarithmic Sobolev inequality* becomes

$$\mathrm{Ent}_\mu(f) \le \frac{1}{2} \int \frac{\|\nabla f\|_2^2}{f} d\mu \tag{7}$$

- **Remark** (***Modified Logarithmic Sobolev Inequality via Convex Cost and Duality***)
  For some ***convex non-negative cost*** $c: \mathbb{R}^n \to \mathbb{R}_+$, ***the convex conjugate*** of c (Legendre transform of c) is defined as

$$c^*(x) := \sup_y \{\langle x\,,\, y\rangle - c(y)\}$$

  Then we can obtain ***the modified logarithmic Sobolev inequality***

$$\mathrm{Ent}_\mu(f) \le \int f^2\, c^*\left(\frac{\nabla f}{f}\right) d\mu \tag{8}$$

2 The Entropy Methods

2.1 Tensorization Property of Φ-Entropy

- **Remark** Recall that the Φ-entropy for $\Phi(x) = x\log(x)$ as

$$H_\Phi(X) = \mathrm{Ent}(X) := \mathbb{E}\left[X \log X\right] - \mathbb{E}\left[X\right] \log\left(\mathbb{E}\left[X\right]\right).$$

  The variational formulation of $H_\Phi(X)$ is

$$\mathrm{Ent}(X) = \sup_T \{X\left(\log(T) - \log(\mathbb{E}\left[T\right])\right)\}$$

-

2.2 Herbst's Argument

2.3 Bounded Difference Inequality

2.4 Modified Logarithmic Sobolev Inequalities

2.5 Concentration of Convex Lipschitz Functions

2.6 Exponential Tail Bounds for Self-Bounding Functions

References

Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press, 2013.