

Log Sobolev Inequality

Reference: ① Ramon van Handel, Probability in High Dimension, Lecture notes, Princeton University, 2016 (RVH book)
② Sinho Chewi, Log-concave sampling, Online notes, Nov. 2024

(I closely follow Chaps. 2, 3 in ①)

Setting as before: $\{P_t\}$ a M.P. semigroup,
with $\mathcal{L}$ the semigroup generator and μ the stat.
measure.

Defn: (Entropy of a function)

$$\text{Ent}_\mu(f) = \mu(f \ln f) - \mu(f) \ln(\mu(f)).$$

i.e., expanding: $= E_{z \sim \mu} [f(z) \ln f(z)]$

$$- E_{z \sim \mu} [f(z)] \ln E_{z \sim \mu} [f(z)]$$

$z \in \mathbb{R}^d$

Note: Suppose ν and μ are measures over $\mathbb{R}^d$
s.t. ν abs. cont. w.r.t μ . Then, let us define

$$f(x) = \frac{d\nu}{d\mu}(x), \text{ the Radon-Nikodym}$$

derivative (equivalently the likelihood ratio if densities exist)

Then, $\text{Ent}_{\mu}(f) = KL(\nu \parallel \mu)$

Why? $KL(\nu \parallel \mu) = \int \ln \left(\frac{\nu(x)}{\mu(x)} \right) \cdot \nu(x) dx.$

$$= \int (\ln f(x)) \cdot \left(\frac{\nu(x)}{\mu(x)} \right) \mu(x) dx$$

$$= E_{z \sim \mu} [f(z) \ln f(z)].$$

Note that $E_{z \sim \mu} [f(z)] = \int \frac{\nu(x)}{\mu(x)} \cdot \cancel{\mu(x)} dx = 1$

$$\therefore \ln E_{z \sim \mu} [f(z)] = 0.$$

$$\begin{aligned} \text{Ent}_\mu(f) &= E_{\mu} [f(z) \ln f(z)] - E_{\mu} [f(z)] \ln E_{\mu} [f(z)] \\ &= K L(v||\mu) \quad \text{when } f = \frac{dv}{d\mu} \end{aligned}$$

Theorem 1 (Thm 3.20 in P & H) (Modified Log-Sobolev Inequality (MLSI)): $\{P_t\}$ a Markov semigroup, with generator $\mathcal{L}$ and stat. measure μ . Then,

$$\textcircled{1} \quad \text{Ent}_\mu(f) \leq c \mathcal{E}(\ln f, f) \quad \forall f \in \mathcal{F}$$

(modified LSI)



$$\textcircled{2} \quad \text{Ent}_\mu(P_t f) \leq e^{-t/c} \text{Ent}_\mu(f)$$

(entropic exponential ergodicity).

Further, if $\text{Ent}_\mu(P_t f) \rightarrow 0$ as $t \rightarrow \infty$ (entropic ergodicity), then:

$$\textcircled{3} \quad \mathcal{E}(\ln P_t f, P_t f) \leq e^{-t/c} \mathcal{E}(\ln f, f)$$

$\forall f \in \mathcal{F}, t \geq 0$

$\Rightarrow$ ①, ② above.

Proof: Recall that:

$$\mu(dP_t f) = \frac{d}{dt} \mu(P_t f) = \frac{d}{dt} \mu(f) = 0.$$

Next, $\text{Ent}_\mu(f) = \mu(f \ln f) - \mu(f) \ln(\mu(f)).$

Thus, $\frac{d}{dt} \text{Ent}_\mu(P_t f) = \mu(dP_t f \ln P_t f) + \underbrace{\mu(dP_t f)}_{=0}$
 $= -\varepsilon(\ln P_t f, P_t f).$

Recall:
 $\varepsilon(f, g) = -\langle f, dg \rangle_\mu$
 $= -\mu(f dg)$

$\hookrightarrow$ (*)

$$\begin{aligned} \frac{d}{dt} P_t f \ln P_t f &= (dP_t f) \ln P_t f \\ &+ \cancel{(P_t f)} \cdot \frac{1}{\cancel{(P_t f)}} \cdot d(P_t f) \end{aligned}$$

3 $\Rightarrow$ 1 :

$$\text{Ent}_\mu(f) = \int_0^\infty \frac{d}{dt} \text{Ent}_\mu(P_t f) dt$$

$$\begin{aligned} &= -\int_0^\infty \varepsilon(\ln P_t f, P_t f) dt \leq \varepsilon(\ln f, f) \int_0^\infty e^{-t/c} dt \\ &= c \varepsilon(\ln f, f) \end{aligned}$$

$$\underline{1 \Rightarrow 2} : \quad \frac{d}{dt} \text{Ent}_\mu(P_t f) \stackrel{\text{from } (*)}{=} -\mathcal{E}(\ln P_t f, P_t f)$$

$$\stackrel{\text{from } (1)}{\leq} -\frac{1}{c} \text{Ent}_\mu(P_t f)$$

$$\underline{2 \Rightarrow 1} : \quad \mathcal{E}(\ln f, f) \stackrel{\text{from } (*)}{=} \lim_{t \rightarrow 0} \frac{\text{Ent}_\mu(f) - \text{Ent}_\mu(P_t f)}{t}$$

$$\stackrel{\text{from } (2)}{\geq} \lim_{t \rightarrow 0} \underbrace{\left(\frac{1 - e^{-t/c}}{t} \right)}_{= c} \text{Ent}_\mu(f)$$

Gaussian MLS1 (Example 3.23 in RVH). (here, $f' = \frac{d}{dx} f(x)$)

1-d case: $\mu = N(0, 1)$

$$dx_t = -x_t dt + \sqrt{2} dB_t \quad (\text{OU process})$$

$$\mathcal{E}(f, g) = \mu(f' g') = \langle f', g' \rangle_\mu$$

$$(P_t f)' = e^{-t} P_t f' \quad (\text{see page 9 in Notes 5}).$$

$\nearrow \frac{\partial}{\partial x}$

We now show: $\text{Ent}_n(f) \leq \frac{1}{2} \mathcal{E}(\ln f, f)$

(MLS1 for Gaussian).

First: $(\ln g)' g' = \frac{|g'|^2}{g}$.

$$\therefore (\ln P_t f)' (P_t f)' = e^{-2t} \frac{|P_t f'|^2}{P_t f} \rightarrow \textcircled{1'}$$

By Cauchy-Schwartz inequality,

$$|P_t f'|^2 = \left| P_t \left(\frac{f'}{\sqrt{f}} \sqrt{f} \right) \right|^2$$

$$\leq P_t \left(\frac{|f'|^2}{f} \right) \cdot P_t f$$

$$= P_t ((\ln f)' f') P_t f \rightarrow \textcircled{2'}$$

$\therefore$ From $\textcircled{1'}$, $\textcircled{2'}$, we have:

$$(\ln P_t f)' (P_t f)' \leq e^{-2t} P_t ((\ln f)' f'). \rightarrow \textcircled{3'}$$

$$\begin{aligned}
\therefore \mathcal{E}(\ln P_t P, P_t f) &= \mu \left((\ln P_t f)' (P_t f)' \right) \\
&\stackrel{(3')}{\leq} \mu \left(e^{-2t} P_t ((\ln f)' f') \right) \\
&= e^{-2t} \underbrace{\mu \left(P_t ((\ln f)' f') \right)}_{= \mu} \\
&= e^{-t} \mu \left((\ln f)' f' \right) \\
&= e^{-t} \mathcal{E}(\ln f, f)
\end{aligned}$$

$$\text{i.e., } \mathcal{E}(\ln P_t f, P_t f) \leq e^{-t} \mathcal{E}(\ln f, f) \rightarrow (4')$$

$\therefore$ From Theorem 1, $(3') \Rightarrow (1)$, we have

$$\text{Ent}_\mu(f) \leq \frac{1}{2} \mathcal{E}(\ln f, f) \quad \square$$

n-dimensional version (proved using tensorization, or directly as above using vector calc.)

Theorem 2 (Thm 3.24 in RVH) (Gaussian MLSI)

$$\text{Ent}_\mu(f(x_1, x_2, \dots, x_d)) \leq \frac{1}{2} E \left[\nabla f(x_1, \dots, x_d) \cdot \nabla f(x_1, \dots, x_d) \right]$$

Remark: The result above implies concentration for Gaussians, as below. See RVH for the proof.

Corollary 3 (Thm 3.25, Gaussian concentration)

$$P\left(f(x_1, \dots, x_d) - E[f(x_1, \dots, x_d)] \geq t\right) \leq e^{-t^2/2\sigma^2},$$

where $x_i \sim N(0, 1)$ iid, $\sigma^2 = \|\nabla f\|_\infty^2$,
and any $t \geq 0$.

The Gaussian Log Sobolev Inequality (LSI):

Recall that the MLSI we discussed so far is

$$\begin{aligned} \text{Ent}_\mu(f) &\triangleq \mu(f \ln f) - \mu(f) \mu(\ln f) \\ &\leq \frac{1}{2} \mathcal{E}(\ln f, f). \end{aligned}$$

We can (in the Gaussian case) rewrite this as:

$$\begin{aligned} \frac{1}{2} \text{Ent}_\mu(f^2) &= \mu(f^2 \ln f) - \mu(f^2) \mu(\ln f) \\ &\leq \mathcal{E}(f, f) \end{aligned}$$

Gaussian LSI

Poincaré: $E[f^2] - (E[f])^2 \leq c \|\nabla f\|_2^2$

LSI: $E[f^2 \ln f] - E[f^2] E[\ln f] \leq c \|\nabla f\|_2^2$

This "extra" $\ln f$ results in stronger concentration results (sub exponential vs. subGaussian concentration); as well as stronger convergence results for M.P. ($\|P_t f - \mu(f)\|_{L^2(\mu)}$ vs. $\text{Ent}_\mu(f)$).

Indeed, we can prove that $\text{MLSI} \Rightarrow \text{PI}$

Lemma 4 (Lemma 3.28 in RvH): Suppose that for $\forall f \in \mathcal{F}$, $\text{Ent}_{\mu}(f) \leq c \mathcal{E}(\ln f, f)$. Then, we have the Poincaré inequality: $\text{Var}_{\mu}(f) \leq 2c \mathcal{E}(f, f)$.

Pf: See RvH.



Finally, analogous to the Poincaré inequality, a Bakry-Emery criterion (an algebraic condition) can be used to assert the MLSI. Below, we state the steps in the result and skip the proofs. (problem 3.19 in RvH).

Setting: P_t a reversible and ergodic M.P. with stat. measure μ and semigroup generator $\mathcal{L}$.

Assume that the Carré du Champ $\Gamma(\cdot, \cdot)$ satisfies:

$$\Gamma(f, \phi \circ g) = \Gamma(f, g) (\phi' \circ g)$$

e.g.: Suppose $\Gamma(f, g) = \nabla f \cdot \nabla g$.

$$\begin{aligned}\Gamma(f, \phi \circ g) &= \nabla f \cdot \nabla (\phi \circ g) \\ &= \nabla f \cdot (\phi' \circ g) \nabla g \\ &= (\phi' \circ g) (\nabla f \cdot \nabla g)\end{aligned}$$

$$\begin{aligned}\nabla \phi \circ g &= \nabla_x \phi(g(x)) \\ &= \phi'(g(x)) \nabla g(x) \\ &= (\phi' \circ g) \nabla g(x).\end{aligned}$$

$$= \Gamma(f, g) (\phi' \circ g).$$

① $\mathcal{E}(\ln P_t f, P_t f)$

$$\leq \mu \left(\frac{\Gamma(f, f)}{f} \right)^{1/2} \mu \left(f \Gamma(P_t \ln P_t f, P_t \ln P_t f) \right)^{1/2}$$

② Bakry - Émery criterion: $c \Gamma_2(f, f) \geq \Gamma(f, f)$



$$\mathcal{E}(\ln P_t f, P_t f) \leq e^{-t/c} \mathcal{E}(\ln f, f)^{1/2} \mu \left(f P_t \Gamma(\ln P_t f, \ln P_t f) \right)^{1/2}$$

$$\textcircled{c} \quad \mathcal{E}(\ln P_t f, P_t f) \leq e^{-t/c} \mathcal{E}(\ln f, f)^{1/2} \mathcal{E}(\ln P_t f, P_t f)^{1/2}$$

i.e.,

Bakry - Émery criterion: $c \Gamma_2(f, f) \geq \Gamma(f, f)$



$$\text{Ent}_\mu(f) \leq \frac{c}{2} \mathcal{E}(\ln f, f) \quad \forall f \in \mathcal{F}.$$

$\textcircled{d}$ Let μ be ρ -strongly log-concave measure over $\mathbb{R}^d$ (same as in the Poincaré notes 5).

$$\text{Then, } \text{Ent}_\mu(f^2) \leq \frac{2}{\rho} \int \|\nabla f\|^2 d\mu$$

Thus, as in the Poincaré Inequality case, we can get convergence in KL that does not have any dependence on dimension d .