HIGH-DIMENSIONAL PROBABILITY

MICHAELMAS TERM 2021

1. Let (X, d, μ) be a metric probability space. Suppose that for any Borel subset A of X with $\mu(A) \geq \frac{1}{2}$ and any $\varepsilon > 0$, we have

$$\mu\{x : \ d(x, A) \leq \varepsilon\} \geq 1 - \alpha(\varepsilon)$$

for some function $\alpha : [0, \infty) \to [0, \infty)$. Prove that if $f : X \to \mathbb{R}$ is an L-Lipschitz function and m_f is a median of f with respect to μ, that is,

$$\min\Big\{\mu\{x : f(x) \geq m_f\}, \mu\{x : f(x) \leq m_f\}\Big\} \geq \frac{1}{2},$$

then

$$\forall\ t > 0, \qquad \mu\Big\{x : \ |f(x) - m_f| \geq t\Big\} \leq 2\alpha(t/L).$$

Hint: Apply the assumption to the sets $\{f \geq m_f\}$ and $\{f \leq m_f\}$.

2. Let $f : \Omega \to \mathbb{R}$ be a measurable function on a probability space (Ω, μ) and assume that there exists a value $a_f \in \mathbb{R}$ such that

$$\forall\ t > 0, \qquad \mu\Big\{x : \ |f(x) - a_f| \geq t\Big\} \leq \beta(t)$$

for some function $\beta : [0, \infty) \to [0, \infty)$. Prove the following concentration inequalities for the function f around its median and mean.

(i) If m_f is a median of f with respect to μ and t_0 is such that $\beta(t_0) < \frac{1}{2}$, then

$$\forall\ t > 0, \qquad \mu\Big\{x : \ |f(x) - m_f| \geq t + t_0\Big\} \leq \beta(t).$$

(ii) If $B \stackrel{\text{def}}{=} \int_0^\infty \beta(s)\, \mathrm{d}s < \infty$, then f is μ-integrable and

$$\forall\ t > 0, \qquad \mu\Big\{x : \ |f(x) - \mathbb{E}_\mu[f]| \geq t + B\Big\} \leq \beta(t).$$

3. The Brunn–Minkowski inequality asserts that for any compact sets A, B in $\mathbb{R}^n$,

$$\mathrm{vol}(A + B)^{\frac{1}{n}} \geq \mathrm{vol}(A)^{\frac{1}{n}} + \mathrm{vol}(B)^{\frac{1}{n}}.$$

In this problem we will present an elementary proof of this inequality. It suffices to assume that each A, B is a disjoint union of a finite number of compact boxes with faces parallel to the coordinate hyperplanes as the general case will follow by approximation. Let N be the total number of boxes involved, that is, if A is a union of N_1 boxes and B is a union of N_2 boxes then $N = N_1 + N_2$. Prove the inequality by induction on N via the following steps.

(i) Prove the base case $N = 2$, that is, the case $A = \prod_{i=1}^n [a_i, b_i]$ and $B = \prod_{i=1}^n [c_i, d_i]$.

(ii) Let $Q_1, \ldots, Q_k$ be pairwise disjoint boxes with faces parallel to the coordinate hyperplanes. Prove that there exists a hyperplane H parallel to a coordinate hyperplane such that if H^+ and H^- are the closed half-spaces determined by H, then there exists $j, j' \in \{1, \ldots, k\}$ such that $Q_j \subset H^+$ and $Q_{j'} \subset H^-$.

(iii) For the inductive step, suppose that A, B are unions of N_1 and N_2 boxes respectively such that $N_1 + N_2 = N + 1$. Choose a hyperplane H which satisfies the conclusion of (ii) for the collection of boxes whose union is A and let $A^+ = A \cap H^+$ and $A^- = A \cap H^-$. Observe that both A^+ and A^-

are unions of at most $N_1 - 1$ boxes. By appropriately translating B, notice that in order to deduce the Brunn–Minkowski inequality, we can assume without loss of generality that

$$\frac{\mathrm{vol}(B^+)}{\mathrm{vol}(B)} = \frac{\mathrm{vol}(A^+)}{\mathrm{vol}(A)} \quad \text{and} \quad \frac{\mathrm{vol}(B^-)}{\mathrm{vol}(B)} = \frac{\mathrm{vol}(A^-)}{\mathrm{vol}(A)}, \tag{$*$}$$

where $B^+ = B \cap H^+$ and $B^- = B \cap H^-$. Use the inclusion

$$A + B \supseteq (A^+ + B^+) \cup (A^- + B^-)$$

along with the inductive hypothesis and $(*)$ to complete the proof.

Deduce from the Brunn–Minkowski inequality that for any compact sets A, B,

$$\forall\ \lambda \in (0,1), \qquad \mathrm{vol}(\lambda A + (1-\lambda)B) \geq \mathrm{vol}(A)^{\lambda}\mathrm{vol}(B)^{1-\lambda}.$$

4. (Borell's lemma) A Borel measure μ on $\mathbb{R}^n$ is log-concave if for every compact subsets A, B of $\mathbb{R}^n$ and $\lambda \in (0,1)$, we have

$$\mu(\lambda A + (1-\lambda)B) \geq \mu(A)^{\lambda}\mu(B)^{1-\lambda}.$$

Prove that if K is an origin-symmetric convex set in $\mathbb{R}^n$, then

$$\forall\ t > 1, \qquad \mu(tK) \geq 1 - \mu(K)\left(\frac{1-\mu(K)}{\mu(K)}\right)^{\frac{t+1}{2}}.$$

Hint: Use the inclusion $\frac{2}{t+1}(\mathbb{R}^n \setminus tK) + \frac{t-1}{t+1}K \subseteq \mathbb{R}^n \setminus K$.

5. Let $X_1, \ldots, X_n$ be independent random vectors with values in a Banach space $(B, \|\cdot\|_B)$. Suppose that these random vectors are bounded in the sense that $\|X_i\|_B \leq C$ a.s. for every $i \in \{1, \ldots, n\}$. Show that

$$\mathrm{Var}\Big[\Big\|\frac{1}{n}\sum_{i=1}^{n} X_i\Big\|_B\Big] \leq \frac{C^2}{n}.$$

6. Let $X_1, \ldots, X_n$ be i.i.d. random variables with values in $[0,1]$. Each X_i represents the size of a package to be shipped. The shipping containers are bins of size 1 (so each bin can hold packages whose sizes sum up to at most 1). Let $B_n = f(X_1, \ldots, X_n)$ be the minimal number of bins needed to store the packages. Note that explicitly computing B_n is a hard combinatorial optimization problem. Prove that

$$\mathrm{Var}[B_n] \leq \frac{n}{4} \qquad \text{and} \qquad \mathbb{E}[B_n] \geq n\mathbb{E}[X_1].$$

7. Let $X_1, \ldots, X_n$ be independent random variables taking values in $[a, b]$. Show that if $f : \mathbb{R}^n \to \mathbb{R}$ is a *convex* function, then

$$\mathrm{Var} f(X_1, \ldots, X_n) \leq (b-a)^2 \mathbb{E}[|\nabla f(X_1, \ldots, X_n)|^2].$$

Hint: If $g : \mathbb{R} \to \mathbb{R}$ is a convex function, then $g(x) - g(y) \geq g'(y)(x-y)$ for all $x, y \in \mathbb{R}$.

8. Consider the probability measure $\mathrm{d}\nu(x) = \frac{1}{2}e^{-|x|}\,\mathrm{d}x$ on $\mathbb{R}$. Find a suitable integration by parts formula for ν and use it to show that if $f : \mathbb{R} \to \mathbb{R}$ is a smooth function, then

$$\mathrm{Var}_{\nu} f \leq 4\int_{\mathbb{R}} [f'(x)]^2\,\mathrm{d}\nu(x).$$

9. Let $X = (X_1, \ldots, X_n) \sim N(0, \Sigma)$ be a centered n-dimensional Gaussian random vector with covariance matrix Σ. Show that

$$\mathrm{Var}\Big[\max_{i \in \{1,\ldots,n\}} X_i\Big] \leq \max_{i \in \{1,\ldots,n\}} \mathrm{Var}[X_i].$$

Hint: Write $X = \Sigma^{1/2} Y$ where $Y \sim N(0, \mathrm{Id}_n)$.

10. Let $(P_t)_{t \geq 0}$ be a reversible Markov semigroup with generator $\mathcal{L}$ and stationary measure μ. The corresponding *carré du champ* is the bilinear operator given by

$$\Gamma(f,g) = \frac{1}{2}\Big\{\mathcal{L}(fg) - f\mathcal{L}g - g\mathcal{L}f\Big\}.$$

a. What is the carré du champ of the Ornstein–Uhlenbeck semigroup?

b. Show that the Dirichlet form satisfies

$$\mathcal{E}(f,g) = \int \Gamma(f,g)\, d\mu.$$

The carré du champ $\Gamma(f,f)$ is interpreted as the square gradient of f.

c. Show that $\Gamma(f,f) \geq 0$. *Hint: Use that $P_t f^2 \geq (P_t f)^2$ and the definition of $\mathcal{L}$.*

d. Prove the Cauchy–Schwarz inequality $\Gamma(f,g)^2 \leq \Gamma(f,f)\Gamma(g,g)$.

e. Prove the identity

$$P_t(f^2) - (P_t f)^2 = 2\int_0^t P_{t-s}\Gamma(P_s f, P_s f)\, ds.$$

Hint: Interpolate along the curve $s \mapsto P_{t-s}(P_s f)^2$.

f. Observe that if an inequality of the form

$$\forall\, s > 0, \qquad \Gamma(P_s f, P_s f) \leq \alpha(s) P_s \Gamma(f,f)$$

holds a.s. for some function $\alpha : (0,\infty) \to (0,\infty)$, then we can derive the *local Poincaré inequality*

$$P_t(f^2) - (P_t f)^2 \leq c(t) P_t \Gamma(f,f), \quad \text{where} \quad c(t) = 2\int_0^t \alpha(s)\, ds.$$

Observe that if $c(t) \to c < \infty$ as $t \to \infty$, then this implies the classical Poincaré inequality for f with constant c.

11. Let $(P_t)_{t \geq 0}$ be a reversible Markov semigroup with generator $\mathcal{L}$ and stationary measure μ. The corresponding Γ_2*-operator* is defined by

$$\Gamma_2(f,g) = \frac{1}{2}\Big\{\mathcal{L}\Gamma(f,g) - \Gamma(f,\mathcal{L}g) - \Gamma(\mathcal{L}f,g)\Big\}.$$

a. What is the Γ_2-operator of the Ornstein–Uhlenbeck semigroup?

b. Prove that the following are equivalent for a fixed $c > 0$:

1. $c\Gamma_2(f,f) \geq \Gamma(f,f)$ for all f (Bakry–Émery criterion).
2. $\Gamma(P_t f, P_t f) \leq e^{-2t/c} P_t \Gamma(f,f)$ for all f and t (local ergodicity).
3. $P_t(f^2) - (P_t f)^2 \leq c(1 - e^{-2t/c}) P_t \Gamma(f,f)$ for all f and t (local Poincaré).

Hint: For 1⇒2 evaluate $\frac{d}{ds} P_{t-s}\Gamma(P_s f, P_s f)$. For 3⇒1, compute the first nonzero term of the Taylor expansion of the local Poincaré inequality at $t = 0$.

c. Consider a measure $d\mu(x) = e^{-W(x)}dx$ on $\mathbb{R}^n$ such that $\mathrm{Hess}\, W(x) \succeq \rho \mathrm{Id}_n$ in the positive semidefinite ordering for some $\rho > 0$ and any $x \in \mathbb{R}^n$. The measure μ is the stationary measure of a Markov process whose semigroup has generator

$$\mathcal{L}_\mu f = \Delta f - \langle \nabla W, \nabla f \rangle.$$

Use the Bakry-Émery criterion to derive the following inequality of Brascamp and Lieb: for any smooth $f : \mathbb{R}^n \to \mathbb{R}$, we have

$$\mathrm{Var}_\mu[f] \le \frac{1}{\rho} \int_{\mathbb{R}^n} |\nabla f(x)|^2 \, \mathrm{d}\mu(x).$$

12. Let $a_1, \dots, a_n \in \mathbb{R}$. Prove that

$$\mathbb{P}\Big\{\Big|\sum_{i=1}^n a_i \varepsilon_i\Big| \ge t\Big\} \le 2e^{-t^2/4\sum_{i=1}^n a_i^2},$$

where $(\varepsilon_1, \dots, \varepsilon_n)$ is uniformly distributed on $\{-1,1\}^n$. Deduce Khintchine's inequality: there exists a universal constant $C \in (0,\infty)$ such that for any $p \ge 2$,

$$\Big(\mathbb{E}\Big|\sum_{i=1}^n a_i \varepsilon_i\Big|^p\Big)^{1/p} \le C\sqrt{p}\sqrt{\sum_{i=1}^n a_i^2}.$$

Hint: Recall that $\mathbb{E}|Y|^p = p\int_0^\infty t^{p-1}\mathbb{P}\{|Y| \ge t\}\,\mathrm{d}t$ *for any random variable* Y.

13. Let $\mathcal{S}_n$ be the symmetric group on n elements equipped with the metric

$$\forall\, \sigma,\tau \in \mathcal{S}_n, \qquad d_{\mathcal{S}_n}(\sigma,\tau) = \frac{1}{n}\#\{i \in \{1,\dots,n\} : \sigma(i) \neq \tau(i)\}$$

and the uniform probability measure $\mathbb{P}$. For $j \in \{0,1,\dots,n\}$, consider the σ-algebra $\mathcal{F}_j$ of subsets of $\mathcal{S}_n$ generated by sets of the form

$$A_{i_1,\dots,i_j} = \Big\{\sigma \in \mathcal{S}_n :\ \sigma(1) = i_1,\ \dots\ ,\sigma(j) = i_j\Big\},$$

where $i_1, \dots, i_j$ are distinct elements of $\{1,\dots,n\}$.

a. Prove that for every atom $A = A_{i_1,\dots,i_j}$ of $\mathcal{F}_j$ and every two atoms $B = A_{i_1,\dots,i_j,r}$, $C = A_{i_1,\dots,i_j,s}$ of $\mathcal{F}_{j+1}$ contained in $\mathcal{F}_j$, there exists a bijection $\phi : B \to C$ such that $d_{\mathcal{S}_n}(b,\phi(b)) \le \frac{2}{n}$ for any $b \in B$.

b. Use part a. and the Azuma–Hoeffding inequality to deduce the following theorem of Maurey: if $f : (\mathcal{S}_n, d_{\mathcal{S}_n}) \to \mathbb{R}$ is a 1-Lipschitz function then

$$\forall\, t \ge 0, \qquad \mathbb{P}\{\sigma :\ f(\sigma) - \mathbb{E}f \ge t\} \le e^{-t^2 n/16}.$$

Hint: Consider the martingale $\{f_j\}_{j=0}^n$ *where* $f_j = \mathbb{E}[f|\mathcal{F}_j]$.

14. A partition $\mathcal{P}$ of a set is a refinement of a partition $\mathcal{Q}$ of the same set if any element $P \in \mathcal{P}$ is contained in some element $Q \in \mathcal{Q}$. We say that a metric space (M, d_M) has length at most ℓ if there exists a sequence of partitions $\{M\} = \mathcal{M}^0, \mathcal{M}^1, \cdots, \mathcal{M}^n = \Big\{\{x\} :\ x \in M\Big\}$ of M such that $\mathcal{M}^i$ is a refinement of $\mathcal{M}^{i-1}$ for every $i \in \{1,\dots,n\}$ and positive numbers $a_1,\dots,a_n$ with $\sum_{i=1}^n a_i^2 \le \ell^2$ for which the following property is satisfied. If $i \in \{1,\dots,n\}$ and $A \in \mathcal{M}^{i-1}$, $B, C \in \mathcal{M}^i$ are such that $B \cup C \subseteq A$, then there exists a bijection $\phi : B \to C$ such that $d_M(b,\phi(b)) \le a_i$ for all $b \in B$.

a. Show that any bounded metric space M has length at most $\mathrm{diam}(M)$.

b. Use the Azuma–Hoeffding inequality to prove the following theorem of Schechtman: if (M, d_M, μ) is a metric probability space with length at most ℓ, then any 1-Lipschitz function $f : (M, d_M) \to \mathbb{R}$ satisfies

$$\forall\, t \ge 0, \qquad \mu\{x :\ F(x) - \mathbb{E}_\mu F \ge t\} \le e^{-t^2/4\ell^2}.$$

15. Prove the following partial converse of Herbst's lemma: if X is a σ^2-subgaussian random variable, then

$$\forall\ \lambda \in \mathbb{R}, \qquad \mathrm{Ent}[e^{\lambda X}] \le 2\lambda^2\sigma^2\mathbb{E}[e^{\lambda X}].$$

Hint: Note that $\mathrm{Ent}[e^{\lambda X}]/\mathbb{E}[e^{\lambda X}] = \mathbb{E}[Z\log Z]$ *for* $Z = e^{\lambda X}/\mathbb{E}[e^{\lambda X}]$*. Now use concavity of the logarithm and that* $\mathbb{E}[e^{\lambda(X-\mathbb{E}X)}] \ge 1$.

16. Let $X_1,\ldots,X_n$ be independent random variables taking values in $[a,b]$. Show that if $f:\mathbb{R}^n \to \mathbb{R}$ is a *convex* function, then

$$\mathrm{Ent}[e^{f(X_1,\ldots,X_n)}] \le (b-a)^2\mathbb{E}[|\nabla f(X_1,\ldots,X_n)|^2 e^{f(X_1,\ldots,X_n)}].$$

Deduce that if f is L-Lipschitz, then

$$\forall\ t\ge 0, \qquad \mathbb{P}\{f(X_1,\ldots,X_n) - \mathbb{E}f(X_1,\ldots,X_n) \ge t\} \le e^{-t^2/4(b-a)^2L^2}.$$

Hint: Recall problem 7.

17. Let $(P_t)_{t\ge 0}$ be a reversible and ergodic Markov semigroup with stationary measure μ and assume that the carré du champ (Problem 10) satisfies the chain rule

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

a. Show that for a positive function f, we have

$$\mathcal{E}(\log P_tf, P_tf)^2 \le \mathbb{E}_\mu\Big[\frac{\Gamma(f,f)}{f}\Big]\mathbb{E}_\mu\Big[f\,\Gamma(P_t\log P_tf, P_t\log P_tf)\Big].$$

Hint: Use reversibility and the Cauchy-Schwarz inequality for $\Gamma(\cdot,\cdot)$.

b. Show that the Bakry–Émery criterion $c\Gamma_2(f,f)\ge\Gamma(f,f)$ for all f implies

$$\mathcal{E}(\log P_tf, P_tf)^2 \le e^{-2t/c}\mathcal{E}(\log f, f)\mathbb{E}_\mu\Big[f\,P_t\Gamma(\log P_tf, \log P_tf)\Big].$$

Hint: Use Problem 11 and the chain rule.

c. Show that the above inequality implies

$$\mathcal{E}(\log P_tf, P_tf) \le e^{-2t/c}\mathcal{E}(\log f, f)$$

and deduce that the Bakry–Émery criterion implies the modified log-Sobolev inequality for all positive functions f,

$$\mathrm{Ent}_\mu[f] \le \frac{c}{2}\mathcal{E}(\log f, f).$$

d. Consider a measure $\mathrm{d}\mu(x) = e^{-W(x)}\,\mathrm{d}x$ on $\mathbb{R}^n$ such that $\mathrm{Hess}\,W(x)\ge \rho\,\mathrm{Id}_n$ in the positive semidefinite ordering for some $\rho>0$ and any $x\in\mathbb{R}^n$. Show that μ satisfies the dimension-free log-Sobolev inequality

$$\mathrm{Ent}_\mu[f^2] \le \frac{2}{\rho}\int_{\mathbb{R}^n}|\nabla f(x)|^2\,\mathrm{d}\mu(x).$$

18. Let $X = (X_1,\ldots,X_n)\sim N(0,\Sigma)$ be a centered n-dimensional Gaussian random vector with covariance matrix Σ.

a. Show that $\max_{i=1,\ldots,n} X_i$ is τ^2-subgaussian, where $\tau^2 = \max_{i=1,\ldots,n}\mathrm{Var}X_i$.
Hint: Recall Problem 9.

b. Prove that the mean and median of $\max_{i=1,\ldots,n}X_i$ satisfy

$$\mathbb{E}\Big[\max_{i=1,\ldots,n}X_i\Big] \le \mathrm{med}\Big[\max_{i=1,\ldots,n}X_i\Big] + \sqrt{2\log 2\tau^2}.$$

Hint: Use part a.

Let $(B, \|\cdot\|_B)$ be a Banach space such that B^* is separable. Then, there exists a countable subset $V \subset B^*$ such that

$$\forall\, x \in B, \qquad \|x\|_B = \sup_{v \in V} v(x).$$

Let X be a centered Gaussian random vector in B, that is, a random vector such that $v(X)$ is a centered Gaussian random variable for any $v \in B^*$. Let

$$\sigma^2 \overset{\text{def}}{=} \max_{v \in V} \mathrm{Var}[v(X)].$$

c. Show that $\sigma^2 < \infty$, $\mathbb{E}\|X\|_B < \infty$ and that $\|X\|_B$ is σ^2-subgaussian.

d. Prove the Landau–Shepp–Marcus–Fernique theorem:

$$\mathbb{E}[e^{\alpha\|X\|_B^2}] < \infty \quad \text{if and only if} \quad \alpha < \frac{1}{2\sigma^2}.$$

Hint: For the only if part, use $\mathbb{E}[e^{\alpha\|X\|_B^2}] \geq \sup_{v \in V} \mathbb{E}[e^{\alpha v(X)^2}]$.

19. Let $(\mathbb{X}, d, \mu)$ be a metric probability space satisfying the T_1-inequality

$$\forall\, \nu \in \mathcal{P}_1(\mathbb{X}), \qquad \mathsf{W}_1(\mu, \nu) \leq \sqrt{2\sigma^2 \mathsf{D}(\nu\|\mu)}.$$

a. For a Borel subset S of $\mathbb{X}$ let μ_S be the restriction of μ on S given by $\mu_S(T) = \frac{\mu(S \cap T)}{\mu(S)}$, where $T \subseteq \mathbb{X}$. Prove that if A, B are disjoint subsets of $\mathbb{X}$, then

$$d(A, B) \leq \mathsf{W}_1(\mu_A, \mu_B) \leq \sqrt{2\sigma^2 \log(1/\mu(A))} + \sqrt{2\sigma^2 \log(1/\mu(B))}.$$

b. Deduce that the T_1-inequality implies geometric concentration: if A is a Borel subset of $\mathbb{X}$ with $\mu(A) \geq \frac{1}{2}$ and

$$A_t = \{x \in \mathbb{X} : \ d(x, y) \leq t \text{ for some } y \in A\},$$

then

$$\forall\, t \geq 0, \qquad \mu(A_t) \geq 1 - 2e^{-t^2/4\sigma^2}.$$

20. Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R}, |\cdot|)$ be two measures on the real line and denote by $F(t) = \mu((-\infty, t])$ and $G(t) = \nu((-\infty, t])$ their cumulative distribution functions.

a. Show that for any smooth function $f : \mathbb{R} \to \mathbb{R}$ with compact support,

$$\int_{\mathbb{R}} f \,\mathrm{d}\mu = -\int_{\mathbb{R}} f'(t) F(t)\,\mathrm{d}t.$$

b. Using part a. deduce that

$$\mathsf{W}_1(\mu, \nu) = \int_{\mathbb{R}} |F(t) - G(t)|\,\mathrm{d}t.$$

c. Construct a coupling $M \in \mathcal{C}(\mu, \nu)$ such that

$$\mathbb{E}_M[|X - Y|] = \int_{\mathbb{R}} |F(t) - G(t)|\,\mathrm{d}t.$$

Hint: If U is uniformly distributed on $[0,1]$, what are the distributions of $F^{-1}(U)$ and $G^{-1}(U)$?

21. Let $\mu_1 \otimes \cdots \otimes \mu_n$ be a product probability measure on Ω^n and ν an arbitrary probability measure on Ω^n. Prove Marton's transportation inequality:

$$\inf_{M \in \mathcal{C}(\mu_1, \otimes \mu_n, \nu)} \sum_{i=1}^{n} \mathbb{P}_M\{X_i \neq Y_i\}^2 \leq \frac{1}{2} D(\nu \| \mu_1 \otimes \cdots \otimes \mu_n),$$

where $(X_1, \ldots, X_n, Y_1, \ldots, Y_n)$ has distribution M.
Hint: Use Pinsker's inequality and tensorization.

22. Let $\mu \in \mathcal{P}_2(\mathbb{R}^n, |\cdot|)$ be a probability measure and $\sigma^2 > 0$. We know that if μ satisfies the modified log-Sobolev inequality

$$\forall\, f : \mathbb{R}^n \to \mathbb{R}, \qquad \mathrm{Ent}_\mu[e^f] \leq \frac{\sigma^2}{2} \mathbb{E}_\mu[|\nabla f|^2 e^f]$$

then μ also satisfies the T_2-inequality

$$\forall\, \nu \in \mathcal{P}_2(\mathbb{R}^n), \qquad \mathsf{W}_2(\mu, \nu) \leq \sqrt{2\sigma^2 D(\nu \| \mu)}.$$

We shall prove a converse of this implication for *convex* functions.

a. Prove that for any function $f : \mathbb{R}^n \to \mathbb{R}$

$$\frac{\mathrm{Ent}_\mu[e^{\lambda f}]}{\mathbb{E}_\mu[e^{\lambda f}]} \leq \lambda \Big\{ \mathbb{E}_\nu[f] - \mathbb{E}_\mu[f] \Big\}, \quad \text{where} \quad \mathrm{d}\nu = \frac{e^{\lambda f}}{\mathbb{E}_\mu[e^{\lambda f}]} \mathrm{d}\mu.$$

b. Prove that if f is assumed to be convex, then

$$\forall\, \lambda \geq 0, \qquad \frac{\mathrm{Ent}_\mu[e^{\lambda f}]}{\mathbb{E}_\mu[e^{\lambda f}]} \leq \lambda \inf_{M \in \mathcal{C}(\mu, \nu)} \mathbb{E}_M[\langle \nabla f(Y), Y - X \rangle]$$

and

$$\forall\, \lambda \leq 0, \qquad \frac{\mathrm{Ent}_\mu[e^{\lambda f}]}{\mathbb{E}_\mu[e^{\lambda f}]} \leq -\lambda \inf_{M \in \mathcal{C}(\mu, \nu)} \mathbb{E}_M[\langle \nabla f(X), X - Y \rangle].$$

c. Conclude that if μ satisfies the T_2-inequality, then

$$\forall\, \lambda \geq 0, \qquad \mathrm{Ent}_\mu[e^{\lambda f}] \leq 2\lambda^2 \sigma^2 \mathbb{E}_\mu[|\nabla f|^2 e^{\lambda f}]$$

and

$$\forall\, \lambda \leq 0, \qquad \mathrm{Ent}_\mu[e^{\lambda f}] \leq 2\lambda^2 \sigma^2 \mathbb{E}_\mu[|\nabla f|^2] \mathbb{E}_\mu[e^{\lambda f}].$$

d. Deduce from Herbst's argument that if $\mathbb{E}_\mu[|\nabla f|^2] \leq 1$, then

$$\forall\, t \geq 0, \qquad \mu\{f - \mathbb{E}_\mu f \leq -t\} \leq e^{-t^2/8\sigma^2}.$$

In particular, this consists of a one-sided refinement of the Gaussian concentration inequality in the class of convex functions.

23. Show that the measure $\mu = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$ on $(\mathbb{R}, |\cdot|)$ does not satisfy the T_2-inequality. Deduce that there does not exist $\sigma^2 > 0$ such that for any $n \in \mathbb{N}$ and any 1-Lipschitz function $f : (\mathbb{R}^n, |\cdot|) \to \mathbb{R}$, the random variable $f(\varepsilon_1, \ldots, \varepsilon_n)$ is σ^2-subgaussian, where $\varepsilon_1, \ldots, \varepsilon_n$ are i.i.d. symmetric Bernoulli variables.

24. Let $\varepsilon_1, \ldots, \varepsilon_n$ be independent symmetric Bernoulli random variables $\mathbb{P}\{\varepsilon_i = \pm 1\} = \frac{1}{2}$ and fix a set $T \subseteq \mathbb{R}^n$. Consider the random variable

$$Z = \sup_{t \in T} \sum_{k=1}^{n} \varepsilon_k t_k.$$

a. Use the bounded differences inequality for the variance to prove that

$$\mathrm{Var}[Z] \le 4 \sup_{t\in T} \sum_{k=1}^{n} t_k^2.$$

b. Denote by

$$\tau^2 = \sum_{k=1}^{n} \sup_{t\in T} t_k^2.$$

Use McDiarmid's inequality to show that

$$\forall\ t \ge 0, \qquad \mathbb{P}\{|Z - \mathbb{E}Z| \ge t\} \le 2e^{-t^2/2\tau^2}.$$

c. Denote by

$$\sigma^2 = 4 \sup_{t\in T} \sum_{k=1}^{n} t_k^2.$$

Use the bounded differences inequality for the entropy to prove that

$$\forall\ t \ge 0, \qquad \mathbb{P}\{Z - \mathbb{E}Z \ge t\} \le e^{-t^2/4\sigma^2}.$$

d. Use the Marton–Talagrand concentration inequality to prove that Z is σ^2-subgaussian.

25. Let $X_1,\dots,X_n$ be independent random variables with values in $\mathbb{X}_1,\dots,\mathbb{X}_n$ respectively. For $c_1,\dots,c_n > 0$, consider the distance d_c on $\mathbb{X}_1 \times \cdots \times \mathbb{X}_n$ given by

$$d_c((x_1,\dots,x_n),(y_1,\dots,y_n)) = \sum_{i=1}^{n} c_i \mathbf{1}_{x_i \ne y_i},$$

where $(x_1,\dots,x_n),(y_1,\dots,y_n) \in \mathbb{X}_1\times\cdots\times\mathbb{X}_n$. Use McDiarmid's inequality to prove that for any measurable subset $A \subseteq \mathbb{X}_1 \times \cdots \times \mathbb{X}_n$, we have

$$\forall\ t \ge 0, \qquad \mathbb{P}\Big\{d_c((X_1,\dots,X_n),A) \ge t\Big\} \le \frac{1}{\mathbb{P}(A)} e^{-t^2/2}.$$

26. Let $(P_t)_{t\ge 0}$ be a reversible Markov semigroup with stationary measure μ and fix $c > 0$. Prove that the log-Sobolev inequality

$$\mathrm{Ent}_\mu[f^2] \le 2c\mathcal{E}(f,f) \quad \text{for all } f$$

implies the modified log-Sobolev inequality

$$\mathrm{Ent}_\mu[f] \le \frac{c}{2}\mathcal{E}(\log f, f) \quad \text{for all nonnegative } f.$$

27. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a smooth function with $\int_{\mathbb{R}^n} f(x)^2\, \mathrm{d}x = 1$ and denote by λ the Lebesgue measure on $\mathbb{R}^n$.

a. Use the Gaussian log-Sobolev inequality to deduce that

$$\mathrm{Ent}_\lambda[f^2] \le 2\int_{\mathbb{R}^n} |\nabla f|^2\, \mathrm{d}x - \frac{n}{2}\log(2\pi) - n.$$

Hint: Consider the function $g = \frac{f^2 e^{|x|^2/2}}{\sqrt{2\pi}}$.

b. Apply the inequality of part a. to $f_\sigma(x) = \sigma^n f(\sigma x)$ for a suitable choice of $\sigma > 0$ to deduce the Euclidean log-Sobolev inequality:

$$\mathrm{Ent}_\lambda[f^2] \le \frac{n}{2}\log\left(\frac{2}{n\pi e}\int_{\mathbb{R}^n} |\nabla f|^2\, \mathrm{d}x\right).$$

28. Let $\mathcal{C}_n = (\{-1,1\}^n, \sigma_n)$ where σ_n is the uniform probability measure on $\{-1,1\}^n$. Recall that the discrete heat flow acts on the Walsh expansion as

$$\forall\ t \geq 0, \qquad P_t\Big(\sum_{S\subseteq\{1,\ldots,n\}} c_S w_S\Big) = \sum_{S\subseteq\{1,\ldots,n\}} e^{-t|S|} c_S w_S.$$

Consider the operator $\Delta^{-1/2}$ whose action on a function on $\mathcal{C}_n$ is given by

$$\Delta^{-1/2}\Big(\sum_{S\subseteq\{1,\ldots,n\}} c_S w_S\Big) = \sum_{\emptyset\neq S\subseteq\{1,\ldots,n\}} \frac{c_S}{|S|^{1/2}} w_S.$$

a. For a function $f : \mathcal{C}_n \to \mathbb{R}$, prove that

$$\mathrm{Var}_{\sigma_n} f = \sum_{i=1}^n \|\Delta^{-1/2}\partial_i f\|_{L_2(\sigma_n)}^2.$$

b. Prove that if $\mathbb{E}_{\sigma_n} g = 0$, then

$$\Delta^{-1/2} g = \frac{1}{\sqrt{\pi}} \int_0^\infty P_t g \,\frac{\mathrm{d}t}{\sqrt{t}}.$$

c. Use hypercontractivity to deduce that there exists an absolute constant $C \in (0,\infty)$ such that if $\mathbb{E}_{\sigma_n} g = 0$, then

$$\|\Delta^{-1/2} g\|_{L_2(\sigma_n)} \leq \frac{C\|g\|_{L_2(\sigma_n)}}{1 + \sqrt{\log(\|g\|_{L_2(\sigma_n)}/\|g\|_{L_1(\sigma_n)})}}.$$

d. Combine the above to derive a proof of Talagrand's influence inequality.

29. Let $\mathrm{d}\nu_n(x) = \frac{1}{2^n} e^{-\|x\|_1}\,\mathrm{d}x$ be the symmetric exponential measure on $\mathbb{R}^n$. For $p \in \{1,2\}$, we will use the notation

$$B_p^n = \Big\{x \in \mathbb{R}^n : \sum_{i=1}^n |x_i|^p \leq 1\Big\}.$$

Prove the following geometric form of Talagrand's two-level concentration inequality for ν_n: if A is a Borel subset of $\mathbb{R}^n$, then

$$\forall\ r > 0, \qquad 1 - \nu_n\Big(A + c_1\sqrt{r}B_2^n + c_2 r B_1^n\Big) \leq \frac{1}{\nu_n(A)} e^{-r},$$

where $c_1, c_2 > 0$ are universal constants.

30. The Gaussian isoperimetric inequality asserts that if $A \subseteq \mathbb{R}^n$ is a measurable set such that $\gamma_n(A) = \gamma_n(H)$ for some half-space H, then $\gamma^+(\partial A) \geq \gamma^+(\partial H)$. Use this statement to deduce the following stronger form of isoperimetry: under the assumptions above, we have $\gamma_n(A_r) \geq \gamma_n(H_r)$ for all $r > 0$, where

$$C_r = \Big\{x \in \mathbb{R}^n : \text{ there exists } y \in C \text{ with } |x - y| \leq r\Big\}.$$

Hint: Assume that A is a finite union of Euclidean balls and differentiate the function $v(r) = \Phi^{-1}(\gamma_n(A_r))$, where Φ is the CDF of the normal distribution.