Exercise Sheet 4

We do not expect you to solve all the exercises in this sheet. You can use these exercises as useful references and perhaps to practice in the future.

Guided Exercise

Exercise 1

In this exercise we will prove Gaussian concentration with a slightly weaker constant, using a general method of the 'τ property' introduced by Maurey.

1. For any function $c : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, define the c-transform of a function $\varphi : \mathbb{R}^n \to \mathbb{R}$ as

   $$\varphi^c(y) = \inf_{x \in \mathbb{R}^n} c(x, y) - \varphi(x).$$

   Check that for $c(x, y) = \|x - y\|^2/2$, $\|y\|^2/2 - \varphi^c(y) = \mathcal{L}(\varphi(x) - \|x\|^2/2)$, where $\mathcal{L}$ is the Legendre transform.

2. We say that a probability space $(X, \mathcal{A}, \mu)$ satisfies **the τ property** with respect to a cost function c if

   $$\int e^{\varphi(x)} d\mu(x) \int e^{\varphi^c(y)} d\mu(y) \leq 1.$$

   Show that if a pair (μ, c) satisfies the τ-property, then for any measureable set $A \subset X$ and any $t > 0$,

   $$1 - \mu\left(\{x : \inf_{x \in A} c(x, y) \leq t\}\right) \leq \mu(A)^{-1} e^{-t}.$$

3. Recall the Prékopa-Leindler inequality from Lesson 2. Use it to show that the cost function $c(x, y) = \|x - y\|^2/4$ and the Gaussian measure satisfy the τ-property.

4. Conclude a concentration result for sets of measure at least $1/2$ in Gaussian space.

Additional Exercises

Exercise 2

Generalize the argument of the guided exercise to show a concentration inequality for a measure μ with density $e^{-u(x)}$ with $u(X)$ convex and satisfying $\nabla^2 u \geq kI$ for some $k > 0$.

Exercise 3

We say that a probability space $(X, \mathcal{A}, \mu)$ satisfies **the weak τ property** with respect to a cost function c if

$$e^{\int \varphi(x) d\mu(x)} \int e^{\varphi^c(y)} d\mu(y) \leq 1.$$

Show that if a pair (μ, c) satisfies the weak τ property, then for any measureable set $A \subset X$ and any $t > 0$,

$$1 - \mu(\{x : \inf_{y \in A} c(x, y) < t\}) \leq e^{-\mu(A)t}$$

Exercise 4

Linearize the following log-Sobolev inequality, i.e. assume that (X, d, μ) is a probability space such that every Lipschitz function satisfies

$$\mathrm{Ent}(f^2) \leq 2\beta \mathbb{E}[|\nabla f|^2],$$

i.e. apply the inequality to the function $1+\varepsilon g$. Show that this gives a Poincaré inequality of the form

$$Var(g) \leq \beta \mathbb{E}(|\nabla g|^2).$$

Exercise 5

1. Follow what was done in class to prove the Gaussian concentration inequality for any M-Lipschitz function: Let $X_1, \ldots, X_n$ be i.i.d. copies of $\mathcal{N}(0,1)$ and let $f : \mathbb{R}^n \to \mathbb{R}$ be M-Lipschitz in the sense that,

   $$|f(x) - f(y)| \leq M|x-y|, \quad x, y \in \mathbb{R}^n,$$

   where $|\cdot|$ on the right-hand side is the Euclidean norm. Then for each $t > 0$,

   $$P(f(X) - Ef(X) > t) \leq \mathrm{e}^{-\frac{t^2}{2M^2}}.$$

2. Denote $f(x) := \max_{i=1,\ldots,n} x_i$. Prove that for any $n \times n$-matrix A,

   $$|f(Ax) - f(Ay)| \leq \sqrt{\max_{i=1,\ldots,n} (A^T A)_{ii}}|x-y|, \quad x, y \in \mathbb{R}^n,$$

   with $|x-y|$ denoting the Euclidean norm of $x-y$ on the right-hand side.

3. Prove the Borell-TIS(Tsirelson–Ibragimov–Sudakov) inequality. Let X be a centered Gaussian on $\mathbb{R}^n$ and set

   $$\sigma_X^2 := \max_{i=1,\ldots,n} E\left(X_i^2\right).$$

   Then for each $t > 0$,

   $$P\left(\left|\max_{i=1,\ldots,n} X_i - E\left(\max_{i=1,\ldots,n} X_i\right)\right| > t\right) \leq 2\mathrm{e}^{-\frac{t^2}{2\sigma_X^2}}.$$

Exercise 6

In this exercise we will repeat Herbst's argument on the discrete hypercube, to obtain the following theorem. For $x \in \{-1,1\}^n$ denote by $\bar{x}^{(i)}$ the vector $(x_1, \ldots, x_{i-1}, -x_i, x_{i+1}, \ldots, x_n)$ and similarly for a random vector X denote $\bar{X}$ the random vector with the i-th coordinate with flipped sign.

1. Let $f : \{-1,1\}^n \to \mathbb{R}$ be an arbitrary real-valued function defined on the n-dimensional binary hypercube and assume that X is uniformly distributed over $\{-1,1\}^n$. Define

   $$\mathcal{E}(f) = \frac{1}{4}\mathbb{E}\left[\sum_{i=1}^n \left(f(X) - f\left(\bar{X}^{(i)}\right)\right)^2\right] = \frac{1}{2}\mathbb{E}\left[\sum_{i=1}^n \left(f(X) - f\left(\bar{X}^{(i)}\right)\right)_+^2\right]$$
   $$\mathrm{Ent}(f) = \mathbb{E}[f(X)\log(f(X))] - \mathbb{E}f(X)\log\mathbb{E}f(X),$$

   Then the Log-Sobolev inequality is that

   $$\mathrm{Ent}\left(f^2\right) \leq 2\mathcal{E}(f).$$

   Write out the statement of the inequality for $g(x) = e^{-\lambda f(x)/2}$.

2. Show that $\mathrm{Ent}\left(g^2\right) \leq \frac{\lambda^2}{4}\sum_{i=1}^n \mathbb{E}\left[\left(f(X) - f\left(\bar{X}^{(i)}\right)\right)_+^2 e^{\lambda f(X)}\right]$.

3. Define $F(\lambda) = \mathbb{E}[e^{\lambda f(X)}]$, and $v = \max_{x\in\{-1,1\}^n} \sum_{i=1}^n \left(f(x) - f\left(\bar{x}^{(i)}\right)\right)_+^2$. Write out the previous inequality in terms of f and F, and solve the differential inequality you get to obtain

   $$F(\lambda) \leq e^{\lambda\mathbb{E}f(X) + \lambda^2 v/4}.$$

4. Use Markov's inequality and optimize over λ to conclude the following concentration result on the discrete hypercube:
   Let $f : \{-1,1\}^n \to \mathbb{R}$ and assume that X is uniformly distributed on $\{-1,1\}^n$. Let $v > 0$ be such that
   $$\sum_{i=1}^{n} \left(f(x) - f\left(\bar{x}^{(i)}\right)\right)_+^2 \leq v$$
   for all $x \in \{-1,1\}^n$. Then the random variable $Z = f(X)$ satisfies, for all $t > 0$,
   $$\boldsymbol{P}\{Z > \mathbb{E}Z + t\} \leq e^{-t^2/v} \quad \text{and} \quad \boldsymbol{P}\{Z < \mathbb{E}Z - t\} \leq e^{-t^2/v}$$

Exercise 7

In this exercise we give another proof of Gaussian concentration with a worse constant.

1. Let X and Y be two $\mathcal{N}(0, I_n)$, independant random vectors, and $f : \mathbb{R}^n \to \mathbb{R}$ 1-Lipschitz. Define
   $$X_t = \cos(t\pi/2)X + \sin(t\pi/2)Y; \quad Y_t = -\sin(t\pi/2)X + \cos(t\pi/2)Y$$
   and show that they are both also $\mathcal{N}(0, I_n)$ and independent.
2. Prove that
   $$f(Y) - f(X) = \int_0^1 \frac{d}{dt} f(X_t)\, dt = \frac{\pi}{2} \int_0^1 Y_t \cdot \nabla f(X_t)\, dt.$$
   and conclude
   $$\mathbb{E}[\exp(\lambda f(Y) - \lambda f(X))] \leq \int_0^1 \mathbb{E}\left[\exp\left(\lambda \frac{\pi}{2} Y_t \cdot \nabla f(X_t)\right)\right] dt.$$
3. Use (2.) and bound the right hand side to obtain
   $$\int_0^1 \mathbb{E}\left[\exp\left(\lambda \frac{\pi}{2} Y_t \cdot \nabla f(X_t)\right)\right] dt \leq \exp\left(\frac{\lambda^2 \pi^2}{8}\right)$$
   (recall that X_t has bounded variance and that f is 1-Lip).
4. Using independence and Jensen's inequality, show that
   $$\mathbb{E}[\exp(\lambda f(Y))] \leq \mathbb{E}[e^{\lambda f(X) + \frac{\lambda^2 \pi^2}{8}}].$$
5. Apply the inequality to two independent copies of X to get a Gaussian concentration inequality. Compare to the inequality you got in class.

Exercise 8

One dimensional Poincaré inequalities. Let $X \sim \mu$ and $Y \sim \nu$ be two measures on $\mathbb{R}$ and let

$$F_\mu(t) := \mathbb{P}(X \leq t) \text{ and } F_\nu(t) := \mathbb{P}(Y \leq t),$$

be their cumulative distribution functions, $F_\mu, F_\nu : \mathbb{R} \to [0,1]$. Assume that both F_ν and F_μ are invertible and let $F_\mu^{-1}, F_\nu^{-1} : [0,1] \to \mathbb{R}$ be the inverse functions. Define $T_\mu^\nu := F_\nu^{-1} \circ F_\mu$.

- Show that $\text{Law}(T_\mu^\nu(X)) = \nu$. That is, for almost every $t \in \mathbb{R}$,
  $$\mathbb{P}\left(T_\mu^\nu(X) \leq t\right) = \mathbb{P}(Y \leq t).$$

Now, suppose that ν is a log-concave measure supported on $[a,b]$ and that μ is the standard Gaussian on $\mathbb{R}$.

- Use the fact that, in the above case, T_μ^ν is $10(b-a)$-Lipschitz to deduce that

$$C_p(\nu) \leq 100(b-a)^2.$$

Suppose instead that $d\mu = \frac{1}{2}e^{-|x|}dx$ is the Laplace distribution and that $d\nu = d \cdot x^{d-1}\mathbf{1}_{[0,1]}dx$. These are the types of distributions which appear in the localization lemma.

- Compute T_μ^ν and deduce an appropriate bound on the Poincaré constant of ν.

Exercise 9

The Brascamp-Lieb inequality. Let γ_d be the standard Gaussian in $\mathbb{R}^d$ and let ν be a measure on $\mathbb{R}^d$ of the form $d\nu = e^{-\varphi}d\gamma_d$, for some convex function φ. We say that ν is *more log-concave than the Gaussian.* In 2000' Caffarelli proved that when ν is more log-concave than Gaussian, there exists a 1-Lipschitz map $T : \mathbb{R}^d \to \mathbb{R}^d$ such that $\text{Law}(T(G)) = \nu$, for $G \sim \gamma_d$.

- Use Caffarelli's theorem to prove that

$$C_p(\nu) \leq 1.$$

Exercise 10

Other uses of the localization lemma. Consider the following result by Carbery and Wright: Let $X \sim \mu$ be a measure on R of the form $d\mu = \ell(t)^{d-1}\mathbf{1}_{[0,1]}dt$, where $\ell : [0,1] \to \mathbb{R}_+$ is a linear function, and let $f : \mathbb{R} \to \mathbb{R}$ be a degree p polynomial. Then, for any $\varepsilon > 0$,

$$\mathbb{P}\left(|f(X)| \leq \varepsilon\right) \leq C\left(\frac{\varepsilon}{\mathbb{E}\left[|f(X)|\right]}\right)^{\frac{1}{d}},$$

for a universal constant $C > 0$.

- Use the above result and the localization lemma to prove the Carbery-Wright anti-concentration theorem: If $K \subset \mathrm{R}^d$ is a convex body and $X \sim \mu_K$ is uniform on K, then, for any degree d polynomial $f : \mathrm{R}^d \to \mathrm{R}$,

$$\mathbb{P}\left(|f(X)| \leq \varepsilon\right) \leq C\left(\frac{\varepsilon}{\mathbb{E}\left[|f(X)|\right]}\right)^{\frac{1}{d}},$$

where C is a universal constant.