401-3382-25L: Log-Sobolev Inequalities and Markov Semigroups

Lecture 3 – Spectral gap, hypercontractivity and transport-ineq

Lecturer: Yuansi Chen *Week 5-6* *Spring 2025*

Key concepts:

- Hermite polynomials
- eigenvectors of OU semigroup
- spectral gap and Poincaré inequality
- hypercontractivity and log-Sobolev inequality
- transport inequalities
  - Pinsker inequality
  - W_1 transport inequality
  - W_2 transport inequality

3.1 Hermite polynomials

We first focus on the one dimensional case $n = 1$.

We can equip $L^2(\gamma_1) = \left\{f : \mathbb{R} \to \mathbb{R} \mid \int f^2(x) d\gamma_1(x) < \infty\right\}$ with an inner product

$$\langle f, g\rangle = \int f g d\gamma_1.$$

The inner product is well-defined because of Cauchy-Schwarz inequality. The set of real polynomials is

$$\mathbb{R}[X] := \left\{\sum_{i=0}^{n} a_i X^i \mid n \in \mathbb{N}, a_i \in \mathbb{R}, \forall i \in \{0, \dots, n\}\right\}.$$

It is clear that $\mathbb{R}[X] \subseteq L^2(\gamma_1)$, because of the fast decay of the Gaussian tail.

Lemma 1. *$\mathbb{R}[X]$ is dense in $L^2(\gamma_1)$.*

Proof. For any $f \in L^2(\gamma_1)$, consider the Laplace transform, for any $t \in \mathbb{R}$.

$$\Phi(t) = \int e^{tx} f(x) \gamma_1(dx).$$

$\Phi(t)$ is finite because by Cauchy-Schwarz inequality

$$(\Phi(t))^2 = \left(\int e^{tx} f(x) \gamma_1(dx) \right)^2 \leq \int f^2 d\gamma_1 \int e^{2tx} \gamma_1(dx) < \infty.$$

Similarly, for any $j \in \mathbb{N}$, the j-th order derivative is also well defined

$$\Phi^{(j)}(t).$$

Additionally, Φ is real analytic on a neighborhood of 0 (one way to see this is to place in $\mathbb{C}$. $\Phi(t)$ is real analytic iff it is restriction of a holomorphic function on $\mathbb{C}$. e^{tx} is holomorphic. Apply Morera's theorem.). If $f \perp \mathbb{R}[X]$ in $L^2(\mathbb{R})$, then the derivatives of any order of Φ is zero. Together with the fact that Φ is analytic, we conclude that $\Phi = 0$. By the uniqueness of Laplace transform, $f = 0$ in $L^2(\gamma_1)$. $\square$

Hermite polynomials The Hermite polynomials $(H_k)_{k \geq 0}$ are the orthogonal polynomials obtained from the canonical basis of $\mathbb{R}[X]$ by using the Gram-Schmidt orthogonalization for the inner product $\langle , \rangle$ in $L^2(\gamma_1)$. They are normalized in such a way that the leading coefficient is always 1. The first few of them are

$$\begin{aligned} H_0(x) &= 1 \\ H_1(x) &= x \\ H_2(x) &= x^2 - 1 \\ H_3(x) &= x^3 - 3x \\ H_4(x) &= x^4 - 6x^2 + 3. \end{aligned}$$

The fact that $\mathbb{R}[X]$ is dense in $L^2(\gamma_1)$ means that $(H_k)_{k \geq 0}$ form a complete orthogonal system in the Hilbert space $L^2(\gamma_1)$.

Lemma 2 (Properties of Hermite polynomials)**.** *Hermite polynomials satisfy*

(a) Generating function: for any $k \geq 0$ and $x \in \mathbb{R}$,

$$H_k(x) = \frac{d^k}{ds^k} G_x(0), \text{ where } G_x(s) = e^{sx - \frac{1}{2}s^2} = \sum_{k=0}^{\infty} \frac{s^k}{k!} H_k(x).$$

(b) Three terms recursion formula: for any $k \geq 0$ and $x \in \mathbb{R}$,

$$H_{k+1}(x) = x H_k(x) - k H_{k-1}(x),$$

with convention $H_{-1} = 0$.

(c) *Recursive differential equation: for any $k \geq 0$ and $x \in \mathbb{R}$,*

$$H_k'(x) = kH_{k-1}(x).$$

(d) *Differential equation: for any $k \geq 0$ and $x \in \mathbb{R}$,*

$$H_k''(x) - xH_k'(x) + kH_k(x) = 0.$$

(e) *The sequence $(H_k/\sqrt{k!})_{k\geq 0}$ is orthonormal in $L^2(\gamma_1)$ and its span is dense.*

Proof. First, we prove (b) the recursion formula by induction. It is verified for $k = 0, 1$. Suppose it is verified up to $k-1$, for k, we need to show that

$$H_{k+1}(x) = xH_k(x) - kH_{k-1}(x)$$

is orthogonal to all lower degree polynomials, as it already has the correct degree and leading coefficient. For $j <= k-2$, $\langle H_{k+1}, H_j \rangle = 0$ follows from the definition and the orthogonality of previous polynomials. For k,

$$\begin{aligned} \langle H_{k+1}, H_k \rangle &= \langle xH_k - kH_{k-1}, H_k \rangle \\ &= \langle xH_k, H_k \rangle \\ &= \int xH_k^2(x) d\gamma_n(x) = 0 \end{aligned}$$

because, γ_n is symmetric around 0. For $k-1$,

$$\begin{aligned} \langle H_{k+1}, H_{k-1} \rangle &= \langle xH_k - kH_{k-1}, H_{k-1} \rangle \\ &= \langle xH_k, H_{k-1} \rangle - k \langle H_{k-1}, H_{k-1} \rangle \\ &\overset{(i)}{=} 0 \end{aligned}$$

(i) follows from integration by parts.

Second, we prove (a). It suffices to show that the polynomials defined via the generating function satisfies the same recursion formula. We differentiate the series term-by-term.

$$\frac{\partial}{\partial s} G(x,s) = \frac{\partial}{\partial s}\left(e^{sx - \frac{s^2}{2}}\right) = (x-s)e^{sx-\frac{s^2}{2}}$$

$$\frac{\partial}{\partial s} G(x,s) = \sum_{k=0}^{\infty} \frac{H_k(x)}{k!} ks^{k-1} = \sum_{k=1}^{\infty} \frac{H_k(x)}{(k-1)!} s^{k-1}$$

We identify the terms in the series.

(c) and (d) follows from (b).

(e) follows from Plancherel's theorem. $\square$

Theorem 3.1.1 (Hermite polynomials as eigenvectors of Ornstein-Uhlenbeck semigroup)**.** *For any $k \geq 0$ and $t \geq 0$, the polynomials H_k is an eigenvector of the OU semigroup P_t (respectively OU infinitesimal generator L) associated to the eigenvalue e^{-kt} (respectively $-k$). In other words, for any $f = \sum_{k\geq 0} a_k H_k \in L^2(\gamma_1)$ with $a_k = \frac{1}{k!}\langle f, H_k\rangle$ and for any $t \geq 0$,*

$$Lf = -\sum_{k\geq 0} k a_k H_k$$

$$P_t f = \sum_{k\geq 0} e^{-kt} a_k H_k.$$

Proof of Thm 3.1.1. Recall that for OU semigroup $L = \Delta - x \cdot \nabla$. According to Lemma 2 (Differential equation), we have

$$LH_k = -kH_k.$$

This verifies that H_k is an eigenvector of L for eigenvalue k.

Recall that

$$P_t(f)(x) = \mathbb{E}[f(X_t) \mid X_0 = x] = \int_{\mathbb{R}^n} f\left(\sqrt{\rho}x + \sqrt{1-\rho}y\right)\gamma_1(dy),$$

where $\rho = e^{-2t}$. Apply it to $x \mapsto G_x(s) = e^{sx-\frac{1}{2}s^2}$, gives

$$\begin{aligned}
P_t(G_\cdot(s)) &= \mathbb{E}\left[e^{s\left(\sqrt{\rho}x+\sqrt{1-\rho}Z\right)-\frac{1}{2}s^2}\right] \\
&= e^{s\left(\sqrt{\rho}x\right)-\frac{1}{2}s^2}\mathbb{E}\left[e^{s\sqrt{1-\rho}Z}\right] \\
&\overset{(i)}{=} e^{s\left(\sqrt{\rho}x\right)-\frac{1}{2}s^2} e^{\frac{1}{2}s^2(1-\rho)} \\
&= G_x(s\sqrt{\rho}).
\end{aligned}$$

(i) follows from completing the square and Gaussian integral.

Using the generating series property of Lemma 2, we have

$$\begin{aligned}
P_t(H_k)(x) &= P_t\left(\frac{d^k}{ds^k}G_\cdot(0)\right)(x) \\
&= \frac{d^k}{ds^k}P_t\left(G_\cdot(s)\right)(x) \mid_{s=0} \\
&= \frac{d^k}{ds^k}G_x(se^{-t}) \mid_{s=0} \\
&= e^{-kt}H_k(x)
\end{aligned}$$

□

We are ready to prove Poincaré inequality via Hermite polynomials

Proof of Gaussian Poincaré inequality via Hermite polynomials. For $f = \sum_{k\geq 0} a_k H_k$, we have

$$a_0 = \langle f, H_0 \rangle = \int f d\gamma_1.$$

and

$$\begin{aligned}
\mathrm{Var}_{\gamma_1}(f) &= \int f^2 d\gamma_1 - a_0^2 \\
&\overset{(i)}{=} \sum_{k\geq 1} k! a_k^2 \\
&\leq \sum_{k\geq 1} k \cdot k! a_k^2 \\
&= -\int f L f d\gamma_1.
\end{aligned}$$

(i) follows from Plancherel's identity. The equality is achieved when $a_k = 0$ for $k \geq 2$, in other words f is a linear function. $\square$

Spectral gap is the length between the first eigenvalue 0 and the second eigenvalue of L. Note that for any symmetric semigroup P_t, we always have

- The first eigenvalue of L is 0, because $P_t \mathbf{1} = \mathbf{1}$ and $L\mathbf{1} = 0$ after taking derivative.
- Eigenvalues of L are negative. Recall the definition of $Lf = \lim_{t\to 0} \frac{P_t f - f}{t}$. So we have

$$\begin{aligned}
\langle f, Lf \rangle &= \lim_{t\to 0} \frac{1}{t} \langle f, (P_t - \mathrm{Id}) f \rangle \\
&= \lim_{t\to 0} \frac{1}{t} \left[\langle f, P_t f \rangle - \langle f, f \rangle \right] \\
&\leq 0
\end{aligned}$$

  The last step follows because P_t is a contraction in $L^2(\mu)$

In the case of OU semigroup, the second eigenvalue is -1, and the gap is 1.

Exponential decay of variance from the spectral gap We have

$$a_0 = \langle f, H_0 \rangle = \int f d\gamma_1 = \int P_t f d\gamma_1.$$

$$\begin{aligned}
\operatorname{Var}_{\gamma_1}(P_t f) &= \int (P_t f - a_0)^2 \, d\gamma_1 \\
&= \sum_{k \geq 1} a_k^2 e^{-2kt} k! \\
&\leq e^{-2t} \sum_{k \geq 1} a_k^2 k! \\
&= e^{-2t} \operatorname{Var}_{\gamma_1}(f).
\end{aligned}$$

Second, we move to multivariate Hermite polynomials Since the multivariate Gaussian is a product measure. We have similar results in the multivariate case.

1. The multivariate real polynomial $\mathbb{R}[X_1, \cdots, X_n]$ is dense in $L^2(\gamma_n)$.
2. The multivariate Hermite polynomials are

   $$H_{k_1,\ldots,k_n}(X_1, \ldots, X_n) = H_{k_1}(X_1) \cdots H_{k_n}(X_n).$$

   They form an orthogonal family in $L^2(\gamma_n)$. Also, they form a complete orthogonal system.
3. The OU infinitesimal generator in $\mathbb{R}^n$ writes

   $$L = \Delta - x \cdot \nabla = L_1 + \cdots + L_n,$$

   where $L_i = \partial_i^2 - x_i \partial_i$ is the one-dimensional OU infinitesimal operator acting on the i-th variable. And

   $$L(H_{k_1,\ldots,k_n}) = -(k_1 + \cdots + k_n) H_{k_1,\ldots,k_n}.$$
4. The spectral gap of L is equal to 1 for any dimension $n \geq 1$. Thus, the dimension-independent Poincaré inequality.

3.2 Hypercontractivity

We have seen that any semigroup P_t is a contraction in L^q for $q \geq 1$. Because of log-Sobolev inequality, the OU semigroup contracts better than that! This is where the "hyper" comes from in hypercontractivity.

Theorem 3.2.1 (Gaussian hypercontractivity (Nelson))**.** *Let $p > 1$ and $t > 0$, and set $p_t = 1 + (p-1)e^{2t}$. Observe that $p_t > p$. Then for any $f \in L^p(\gamma_n)$, we have*

$$\|P_t f\|_{p_t} \leq \|f\|_p .$$

In words, P_t is a bounded operator from $L^p(\gamma_n)$ to $L^{p_t}(\gamma_n)$ and has norm 1. Moreover, if $q > p_t$, then P_t is not even bounded from $L^p(\gamma_n)$ to $L^q(\gamma_n)$.

Proof. Without loss of generality, we may assume $f \geq 0$ since $|P_t f| \leq P_t |f|$. Set $\alpha(t) = \log \|P_t\|_{p_t}$. One want to calculate $\alpha'(t)$, show that $\alpha'(t) \leq 0$ via log-Sobolev inequality and then conclude.

We have

$$\begin{aligned}
\alpha'(t) &= \left(\frac{1}{p_t} \log \int (P_t f)^{p_t} d\gamma_n\right)' \\
&= -\frac{p_t'}{p_t^2} \log \int (P_t f)^{p_t} d\gamma_n + \frac{1}{p_t} \frac{(\int (P_t f)^{p_t} d\gamma_n)'}{\int (P_t f)^{p_t} d\gamma_n} \\
&= -\frac{p_t'}{p_t^2} \log \int (P_t f)^{p_t} d\gamma_n + \frac{1}{p_t} \frac{(\int (p_t' \log P_t f + p_t \frac{LP_t f}{P_t f})(P_t f)^{p_t} d\gamma_n)}{\int (P_t f)^{p_t} d\gamma_n} \\
&= \frac{p_t'}{p_t^2} \frac{1}{\int (P_t f)^{p_t} d\gamma_n} \left[\mathrm{Ent}_{\gamma_n}[(P_t f)^{p_t}] + \frac{p_t^2}{p_t'} \int (LP_t f)(P_t f)^{p_t - 1} d\gamma_n\right].
\end{aligned}$$

From integration by parts, we know that

$$\begin{aligned}
\int \frac{|\nabla (P_t f)^{p_t}|^2}{(P_t f)^{p_t}} d\gamma_n &= p_t^2 \int |\nabla (P_t f)|^2 (P_t f)^{p_t - 2} d\gamma_n \\
&= \frac{p_t^2}{p_t - 1} \left\langle \nabla (P_t f), \nabla (P_t f)^{p_t - 1} \right\rangle \\
&= -\frac{p_t^2}{p_t - 1} \int (LP_t f)(P_t f)^{p_t - 1} d\gamma_n,
\end{aligned}$$

where the last step applies integration by parts. Note that $2(p_t - 1) = p'(t)$, we obtain that

$$\alpha'(t) = \frac{p_t'}{p_t^2} \frac{1}{\int (P_t f)^{p_t} d\gamma_n} \left[\mathrm{Ent}_{\gamma_n}[(P_t f)^{p_t}] - \frac{1}{2} \int \frac{\nabla (P_t f)^{p_t}}{(P_t f)^{p_t}} d\gamma_n\right].$$

By log-Sobolev inequality, we obtain

$$\alpha'(t) \leq 0, \forall t \geq 0.$$

Hence,

$$\log \|P_t f\|_{p_t} = \alpha(t) \leq \alpha(0) = \log \|f\|_p .$$

Finally, if $q > p_t$, we know that $f_\lambda(x) = e^{\langle \lambda, x\rangle}$ achieves equality in log-Sobolev inequality. We have

$$\begin{aligned} \|f_\lambda\|_p &= e^{\frac{1}{2}p|\lambda|^2} \\ P_t f_\lambda &= e^{\frac{1}{2}|\lambda|^2(1-e^{-2t})} f_{\lambda e^{-t}}, \end{aligned}$$

therefore,

$$\frac{\|P_t f_\lambda\|_q}{\|f_\lambda\|_p} = e^{\frac{1}{2}\lambda^2(e^{-2t}(q-1)+1-p)}$$

If $q > p_t = 1 + (p-1)e^{2t}$, then the above quantity tends to infinity as $|\lambda| \to \infty$.

$\square$

Remark 1 (Hypercontractivity vs. log-Sobolev inequality)**.** *This proof shows more generally that a semigroup satisfying the log-Sobolev inequality is hypercontractive. Moreover, it is pretty clear from the argument that the implication can be reversed and that log-Sobolev and hypercontractivity are equivalent. This equivalence between hypercontractivity and log-Sobolev inequality is due to Leonard Gross [Gro75].*

3.3 Transport inequalities

For any $p \geq 1$, let $\mathcal{P}_p(\mathbb{R}^n)$ be the set of probability measures on $\mathbb{R}^n$ with finite moment of order p. In other words, $\mu \in \mathcal{P}_p(\mathbb{R}^n)$ if $|\cdot|^p \in \mathbb{L}^1(\mu)$.

Wasserstein-Kantorovich distance on $\mathcal{P}_p(\mathbb{R}^n)$ It is defined for any $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^n)$ by

$$W_p(\mu,\nu) = \inf_{\pi \in \Pi(\mu,\nu)} \left[\int\int_{\mathbb{R}^n \times \mathbb{R}^n} |x-y|^p \, d\pi(x,y) \right]^{\frac{1}{p}},$$

where $\Pi(\mu,\nu)$ is the set of probability measures on the product space $\mathbb{R}^n \times \mathbb{R}^n$ with marginals μ and ν. It can be shown that W_p is a distance on $\mathcal{P}_p(\mathbb{R}^n)$, and that for any $(\mu_k)_{k\in\mathbb{N}}$ and $\mu \in cP_p(\mathbb{R}^n)$, we have $W_p(\mu_k,\mu) \to 0$ if and only if $\mu_k \to \mu$ with respect to continuous test functions $f : \mathbb{R}^n \to \mathbb{R}$ such that $x \mapsto f(x)/(1+|x|^p)$ is bounded.

Note that by Jensen's inequality, $W_p \leq W_q$ for $p < q$.

Variational definition of Wasserstein distance The Kantorovich-Rubinstein duality (see Section 9.1 [BGL13]) states that for any $p \geq 0$ and $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^n)$,

$$W_p(\mu,\nu)^p = \sup \left[\int f d\mu - \int g d\nu \right],$$

where the supremum runs over all bounded continuous functions $f, g : \mathbb{R}^n \to \mathbb{R}$ such that $f(x) - g(y) \leq |x - y|^p$. For W_1, it can also be written as follows (when the cost function is a distance function)

$$W_1(\mu, \nu) = \sup \left[\int f d\mu - \int f d\nu \right]$$

where the supremum runs over all 1-Lipschitz functions (with respect to $|\cdot|$) functions f.

In general, transport inequalities (or transport cost inequalities) are inequalities compare a transport cost distance (Wasserstein distance) to a fluctuation distance expressed by KL divergence. For example

$$W_p(\mu, \nu) \leq \sqrt{C \mathrm{KL}(\nu \parallel \mu)}.$$

In the following, we discuss three cases

- Pinsker inequality where distance is taken as $\mathbf{1}_{x \neq y}$, and $p = 1$
- $p = 1$, and distance $|\cdot|$
- $p = 2$

3.3.1 Pinsker-Csizsár-Kullback inequality

The first example of a transport inequality is the Pinsker-Csizsár-Kullback inequality.

Theorem 3.3.1 (Pinsker-Csizsár-Kullback inequality)**.**

$$d_{TV}(\mu, \nu) \leq \sqrt{\frac{KL(\nu \parallel \mu)}{2}}.$$

where $d_{TV}(\mu, \nu) = \sup_{A \in \mathcal{B}(\mathbb{R}^n)} |\mu(A) - \nu(A)|$

Proof. Let $f \geq 0$ denotes the Radon-Nikodym derivative $\frac{d\nu}{d\mu}$. Recall the density formula of the TV distance, we have

$$d_{\mathrm{TV}}(\mu, \nu) = \frac{1}{2} \int \left| 1 - \frac{d\nu}{d\mu} \right| d\mu.$$

It suffices to show that

$$\left[\int |1 - f| \, d\mu \right]^2 \leq 2 \mathrm{Ent}_\mu(f).$$

Set $f_s = sf + (1-s)\mathbf{1}, s \in [0,1]$, consider

$$\begin{aligned}\Lambda(s) &:= 2\mathrm{Ent}_\mu(f_s) - \left[\int |1-f_s|\, d\mu\right]^2 \\ &= 2\mathrm{Ent}_\mu(f_s) - s^2 \left[\int |1-f|\, d\mu\right]^2.\end{aligned}$$

We have, since $\mathbb{E}_\mu f_s = 1$

$$\Lambda'(s) = 2\left\{\mathbb{E}\left[(f-1)\log f_s + (f-1)\right] - \left[\mathbb{E}(f-1)\log \mathbb{E} f_s + \mathbb{E}(f-1)\right]\right\} - 2s\left[\mathbb{E}_\mu|1-f|\right]^2$$

$$\Lambda''(s) = 2\mathbb{E}_\mu \frac{(f-1)^2}{f_s} - 2\left[\mathbb{E}_\mu|1-f|\right]^2$$

Note that $\Lambda(0) = \Lambda'(0) = 0$, and by Cauchy-Schwarz inequality, we have

$$\Lambda'' \geq 0.$$

Integrating, we obtain that $\Lambda(s) \geq 0$. In particular, $\Lambda(1) \geq 0$, which is desired. □

3.3.2 Transport inequality W_1 and sub-Gaussian concentration of Lipschitz functions

Bobkov and Götze [BG99] discovered that W_1 transport inequality is the dual reformulation of the sub-Gaussian concentration for Lipschitz functions. The proof relies on the Kantorovich-Rubinstein dual representation of W_1.

Theorem 3.3.2. *For any $\mu \in \mathcal{P}_1(\mathbb{R}^n)$ and any constant $c > 0$, the following two statements are equivalent*

*(a). **Sub-Gaussian bound on Laplace transform of Lipschitz functions:** for any Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}$ and any $\theta \in \mathbb{R}$,*

$$L(\theta) := \int \exp(\theta f) d\mu \leq \exp\left(\frac{c}{2}\theta^2 \|f\|_{Lip}^2 + \theta \int f d\mu\right).$$

*(b). **Transport inequality W_1 for μ:** for any $\nu \in \mathcal{P}_1(\mathbb{R}^n)$,*

$$W_1(\mu, \nu) \leq \sqrt{2cKL(\nu \parallel \mu)}.$$

In particular, both of the above holds for $\mu = \gamma_n$ with $c = 1$ and any $n \geq 1$.

Before we prove the theorem, we recall the variational formula for the entropy.

Lemma 3 (Variational formula for the entropy)**.** *For any measurable function $h : \mathbb{R}^n \to \mathbb{R}_+$ such that $\mathbb{E}_\mu h = 1$, we have*

$$Ent_\mu(h) = \sup_{g \text{ measurable}} \left\{ \int ghd\mu - \log \int e^g d\mu \right\}$$

In words, the entropy is the Legendre transform (convex dual) of the log-Laplace transform.

Proof of Lemma 3. By Jensen's inequality, we have

$$\int \frac{e^g}{h} hd\mu \geq e^{\int ghd\mu - \int (\log h)hd\mu},$$

which gives

$$\log \int e^g d\mu \geq \int ghd\mu - \int h \log hd\mu.$$

Hence,

$$\int ghd\mu - \log \int e^g d\mu \leq \int h \log hd\mu.$$

Taking sup over all g, and note that equality is achieved when $g = \log h$, we conclude. □

Proof of Theorem 3.3.2. Without loss of generality, it suffices to deal with $\mu(f) = 0$ and $\|f\|_{\mathrm{Lip}} = 1$ by translation and scaling.

$(a) \implies (b)$: take $g = \theta f - \frac{c}{2}\theta^2$, using the varitional formula for the entropy, we have

$$\begin{aligned} \mathrm{Ent}_\mu(h) &\geq \int \left(\theta f - \frac{c}{2}\theta^2\right) hd\mu - \log \int e^{\theta f - \frac{c}{2}\theta^2} d\mu \\ &\geq \int \left(\theta f - \frac{c}{2}\theta^2\right) hd\mu. \end{aligned}$$

Then with $\mathbb{E}_\mu h = 1$, we have

$$\int (fh - f)d\mu \leq \frac{c}{2}\theta + \frac{1}{\theta} \int h \log hd\mu.$$

Taking infimum over $\theta > 0$, we obtain

$$\int (fh - f)d\mu \leq \sqrt{2c \int h \log hd\mu}.$$

Taking supremum over f, we conclude using Kantorovich–Rubinstein dual formulation.

$(b) \implies (a)$: note that the argument can be reversed. □

3.3.3 Transport inequality W_2

The following theorem provides the dual reformulation of the W_2 transportation inequality, via infimum convolution. **TODO**

3.4 Exercises

Exercise 1. *For Hermite polynomials,*

- *Show that* $H_5 = x^5 - 10x^3 + 15x$ *via three terms recursion formula.*
- *Show the same via Gram-Schmidt orthogonalization.*
- *Show the same via the generating function.*
- *Compute the expansion of* $x^5 - x$ *under the Hermite polynomials*

Exercise 2. *Derive (c) and (d) in Lemma 2.*

Exercise 3. *Show that the zeros of* $H_n(x)$ *are real and distinct (hint: use recursive differential equation).*

Exercise 4. *Let* P_t *be the OU semigroup. Take* $f(x) = x^2$*. Compute* $P_t f(x)$ *explicitly. Show that*

$$\|P_t f\|_{q_t} \leq \|f\|_2 \,,$$

where $q_t = 1 + (p-1)e^{-2t}$*. Can we take a larger* q_t*?*

Exercise 5. *Log-Sobolev Inequality for general product space domains.* **TODO**

Bibliography

[BG99] Sergej G Bobkov and Friedrich Götze. Exponential integrability and transportation cost related to logarithmic sobolev inequalities. *Journal of Functional Analysis*, 163(1):1–28, 1999.

[BGL13] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348. Springer Science & Business Media, 2013.

[Gro75] Leonard Gross. Logarithmic sobolev inequalities. *American Journal of Mathematics*, 97(4):1061–1083, 1975.