Duality for optimal couplings in free probability

Kyeongsik Nam (joint work with Wilfrid Gangbo, David Jekel, and Dimitri Shlyakhtenko)

University of California, Los Angeles

Berkeley Free Probabilistic Operator Algebras Seminar
June 30, 2021

Non-commutative probability spaces

A tracial W^*-algebra is a pair $\mathcal{A} = (A, \tau)$ with W^*-algebra A and a linear functional $\tau : A \to \mathbb{C}$ (trace) such that

- Unital: $\tau(1) = 1$,
- Positive: $\tau(a^* a) \geq 0$,
- Tracial: $\tau(ab) = \tau(ba)$,
- Faithful: $\tau(a^* a) = 0 \implies a = 0$.

Non-commutative probability spaces

A tracial W^*-algebra is a pair $\mathcal{A} = (A, \tau)$ with W^*-algebra A and a linear functional $\tau : A \to \mathbb{C}$ (trace) such that

- Unital: $\tau(1) = 1$,
- Positive: $\tau(a^* a) \geq 0$,
- Tracial: $\tau(ab) = \tau(ba)$,
- Faithful: $\tau(a^* a) = 0 \implies a = 0$.

Every commutative tracial W^*-algebra is isomorphic to some $L^\infty(\Omega, P)$ with a trace $\tau(X) = \int X \, dP$.

Hence, a tracial W^*-algebra may be viewed as a non-commutative probability space.

Non-commutative probability spaces

classical	non-commutative
$L^\infty(\Omega, P)$	A
expectation $\mathbb{E}$	trace τ
bounded random variable	$X \in A$
bounded real random variable	$X \in A_{\mathrm{sa}}$ (self-adjoint)
bounded $\mathbb{R}^m$-valued random variable	$X = (X_1, \ldots, X_m) \in A_{\mathrm{sa}}^m$

Non-commutative probability spaces

classical	non-commutative
$L^\infty(\Omega, P)$	A
expectation $\mathbb{E}$	trace τ
bounded random variable	$X \in A$
bounded real random variable	$X \in A_{\mathrm{sa}}$ (self-adjoint)
bounded $\mathbb{R}^m$-valued random variable	$X = (X_1, \ldots, X_m) \in A_{\mathrm{sa}}^m$

Let $L^2(A, \tau)$ be the GNS construction from (A, τ):

$$\langle a, b \rangle := \tau(a^* b).$$

Non-commutative laws

- Let $\Sigma_{m,R}$ be the collection of unital, positive, tracial linear maps $\mu : \mathbb{C}\langle X_1, \dots, X_m \rangle \to \mathbb{C}$ satisfying

$$|\mu(X_{i_1} \dots X_{i_n})| \leq R^n.$$

Each element in $\Sigma_{m,R}$ is called non-commutative law.
It encodes the non-commutative moments of some tuple of non-commutative random variables.

Non-commutative laws

- Let $\Sigma_{m,R}$ be the collection of unital, positive, tracial linear maps $\mu : \mathbb{C}\langle X_1, \ldots, X_m \rangle \to \mathbb{C}$ satisfying

$$|\mu(X_{i_1} \ldots X_{i_n})| \leq R^n.$$

  Each element in $\Sigma_{m,R}$ is called non-commutative law.
  It encodes the non-commutative moments of some tuple of non-commutative random variables.

- For $X = (X_1, \cdots, X_m) \in A_{\text{sa}}^m$, the associated non-commutative law is

$$\lambda_X(p) = \tau(p(X)).$$

- Tracial non-commutative laws $\leftrightarrow$ tracial W^*-algebras with a specified generating m-tuple up to isomorphism.

Goal of talk

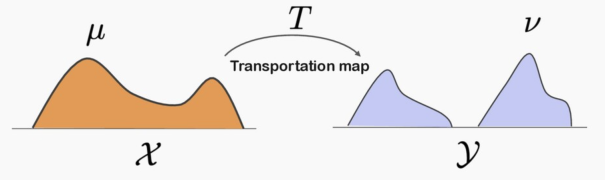
Optimal couplings of two non-commutative laws?

- Review on classical optimal transport theory
- Motivations of optimal couplings in non-commutative setting
- Duality for optimal couplings
- Geodesics via Hopf-Lax semigroup

Classical transportation theory

Monge's optimal transport problem: given $\mu, \nu \in \mathcal{P}(\mathbb{R}^m)$,

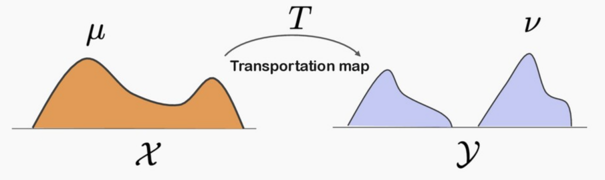
$$\min_{T_\# \mu = \nu} \int_{\mathbb{R}^m} |x - T(x)|^2 d\mu(x).$$



Classical transportation theory

Monge's optimal transport problem: given $\mu, \nu \in \mathcal{P}(\mathbb{R}^m)$,

$$\min_{T_{\#}\mu=\nu} \int_{\mathbb{R}^m} |x - T(x)|^2 d\mu(x).$$



Monge problem can be ill-posed since

- Transport T may not exist.
- Constraint $T_{\#}\mu = \nu$ is not sequentially closed w.r.t. any reasonable weak topology.

Classical optimal couplings

Relaxation of Monge problem? Kantorovich (or Wasserstein) distance

$$d_W^2(\mu,\nu) := \inf_{(X,Y)\text{ coupling}} \mathbb{E}|X-Y|^2.$$

A coupling of μ and ν is a pair of random variables (X,Y) in some probability space such that $X \overset{\text{law}}{\sim} \mu$ and $Y \overset{\text{law}}{\sim} \nu$.

Classical optimal couplings

Relaxation of Monge problem? Kantorovich (or Wasserstein) distance

$$d_W^2(\mu,\nu) := \inf_{(X,Y)\text{ coupling}} \mathbb{E}|X-Y|^2.$$

A coupling of μ and ν is a pair of random variables (X, Y) in some probability space such that $X \overset{\text{law}}{\sim} \mu$ and $Y \overset{\text{law}}{\sim} \nu$.

Properties:

- Set of coupling is compact w.r.t. weak topology in $\mathcal{P}(\mathbb{R}^m \times \mathbb{R}^m)$.
- Minimizer (*optimal coupling*) always exists.
- If $\mu \in \mathcal{P}_2(\mathbb{R}^m)$ is absolutely continuous, then optimal coupling is induced by a gradient of convex function.
- d_W metrizes the weak topology on $\mathcal{P}_2(\mathbb{R}^m)$.

Non-commutative optimal couplings (Biane-Voiculescu)

Recall: Let $\Sigma_{m,R}$ be a collection of unital, positive, tracial linear maps $\mu : \mathbb{C}\langle X_1, \ldots, X_m \rangle \to \mathbb{C}$ satisfying $|\mu(X_{i_1} \ldots X_{i_n})| \leq R^n$.

Non-commutative optimal couplings (Biane-Voiculescu)

Recall: Let $\Sigma_{m,R}$ be a collection of unital, positive, tracial linear maps $\mu : \mathbb{C}\langle X_1, \ldots, X_m \rangle \to \mathbb{C}$ satisfying $|\mu(X_{i_1} \ldots X_{i_n})| \leq R^n$.

- A coupling of μ and ν is a tracial W^*-algebra $\mathcal{A} = (A, \tau)$ and random variables $X, Y \in A_{\text{sa}}^m$ with $\lambda_X = \mu$ and $\lambda_Y = \nu$.
- Wasserstein distance

$$d_W(\mu, \nu) := \inf_{(X,Y) \text{ coupling}} \|X - Y\|_{L^2(A,\tau)_{\text{sa}}^m}$$

Optimal couplings exist by compactness of $\Sigma_{2m,R}$.

Non-commutative optimal couplings (Biane-Voiculescu)

Recall: Let $\Sigma_{m,R}$ be a collection of unital, positive, tracial linear maps $\mu : \mathbb{C}\langle X_1, \ldots, X_m \rangle \to \mathbb{C}$ satisfying $|\mu(X_{i_1} \ldots X_{i_n})| \leq R^n$.

- A coupling of μ and ν is a tracial W^*-algebra $\mathcal{A} = (A, \tau)$ and random variables $X, Y \in A_{\mathrm{sa}}^m$ with $\lambda_X = \mu$ and $\lambda_Y = \nu$.
- Wasserstein distance

$$d_W(\mu, \nu) := \inf_{(X,Y) \text{ coupling}} \|X - Y\|_{L^2(A,\tau)_{\mathrm{sa}}^m}$$

  Optimal couplings exist by compactness of $\Sigma_{2m,R}$.
- d_W defines a metric on $\Sigma_{m,R}$. It does not metrize the weak-$*$ topology (explained in David's talk).

Motivation

Transportation in non-commutative probability spaces.

Theorem (Guionnet-Shlyakhtenko, 14')

Consider a free Gibbs law μ_V with a log-density $V = \frac{1}{2}\sum_{i=1}^{m} X_i^2 + W$. For small enough W, there exists a transport map T from the law of semicircle family to μ_V.

Motivation

Transportation in non-commutative probability spaces.

Theorem (Guionnet-Shlyakhtenko, 14')

Consider a free Gibbs law μ_V with a log-density $V = \frac{1}{2}\sum_{i=1}^{m} X_i^2 + W$. For small enough W, there exists a transport map T from the law of semicircle family to μ_V.

Theorem (Jekel-Li-Shlyakhtenko, 21')

Let μ be any non-commutative law. If $\nabla\phi$ is close to the identity map, then $\nabla\phi$ is an optimal transport from μ to $(\nabla\phi)_{\#}\mu$. In particular, transport map T above is optimal for W small enough.

Systematic theory for optimal transport in non-commutative probability spaces?

Classical Monge-Kantorovich duality

Given $X \overset{\text{law}}{\sim} \mu$ and $Y \overset{\text{law}}{\sim} \nu$,

$$\text{minimizing } \mathbb{E}|X - Y|^2 \iff \text{maximixing } \mathbb{E}\langle X, Y\rangle.$$

Classical Monge-Kantorovich duality

Given $X \overset{\text{law}}{\sim} \mu$ and $Y \overset{\text{law}}{\sim} \nu$,

$$\text{minimizing } \mathbb{E}|X - Y|^2 \iff \text{maximizing } \mathbb{E}\langle X, Y\rangle.$$

Classical duality

$$\sup_{(X,Y) \text{ coupling}} \mathbb{E}\langle X, Y\rangle = \inf \Big(\int f \, d\mu + \int g \, d\nu \Big),$$

where inf is taken over pairs (f, g) of convex functions f, g satisfying $f(x) + g(y) \geq \langle x, y\rangle$.

Classical Monge-Kantorovich duality

Classical duality

Let (X, Y) be a coupling of μ and ν. Then, the following are equivalent.

1. (X, Y) is an optimal coupling.
2. There exist convex functions f, g with $f(x) + g(y) \geq \langle x, y \rangle$ everywhere such that

$$\mathbb{E}f(X) + \mathbb{E}g(Y) = \mathbb{E}\langle X, Y \rangle.$$

3. There exists a convex function f such that $Y \in \partial f(X)$.

For a convex function $f : H \to \mathbb{R}$, we say $y \in \partial f(x)$ (subgradient) if

$$f(x') - f(x) \geq \langle x' - x, y \rangle, \qquad \forall x' \in H.$$

Non-commutative case: tracial W^*-function

- In the classical setting, $(B(0,1), \mathrm{Leb}_{B(0,1)})$ is a universal probability space. In fact, any $\mu \in \mathcal{P}_2(\mathbb{R}^m)$ is a push-forward of $\mathrm{Leb}_{B(0,1)}$ under the gradient of convex function.

Non-commutative case: tracial W^*-function

- In the classical setting, $(B(0,1), \mathrm{Leb}_{B(0,1)})$ is a universal probability space. In fact, any $\mu \in \mathcal{P}_2(\mathbb{R}^m)$ is a push-forward of $\mathrm{Leb}_{B(0,1)}$ under the gradient of convex function.
- In the non-commutative setting, there is no separable tracial W^*-algebra that contains a copy of all others (Ozawa, 04'). Need to consider functions f defined for *all* tracial W^*-algebras.

Non-commutative case: tracial W^*-function

- In the classical setting, $(B(0,1), \mathrm{Leb}_{B(0,1)})$ is a universal probability space. In fact, any $\mu \in \mathcal{P}_2(\mathbb{R}^m)$ is a push-forward of $\mathrm{Leb}_{B(0,1)}$ under the gradient of convex function.
- In the non-commutative setting, there is no separable tracial W^*-algebra that contains a copy of all others (Ozawa, 04'). Need to consider functions f defined for *all* tracial W^*-algebras.

Definition

A tracial W^*-function is a collection $f^{\mathcal{A}} : L^2(\mathcal{A})^m_{\mathrm{sa}} \to (-\infty, \infty]$ of functions for each tracial W^*-algebra such that if $\iota : \mathcal{A} \to \mathcal{B}$ is a trace-preserving unital $*$-homomorphism, then $f^{\mathcal{A}} = f^{\mathcal{B}} \circ \iota$.

Convex functions

Convex functions on non-commutative probability spaces?

In the classical case, consider functions of random variables.
Let $f : \mathbb{R}^m \to \mathbb{R}$ be a convex function and $(\Omega, \mathcal{F}, P)$ be a probability space.
Define $\tilde{f}(X) = \mathbb{E}f(X)$ for random variables $X \in L^2(\Omega, P; \mathbb{R}^m)$.

Convex functions

Convex functions on non-commutative probability spaces?

In the classical case, consider functions of random variables.
Let $f : \mathbb{R}^m \to \mathbb{R}$ be a convex function and $(\Omega, \mathcal{F}, P)$ be a probability space.
Define $\tilde{f}(X) = \mathbb{E}f(X)$ for random variables $X \in L^2(\Omega, P; \mathbb{R}^m)$.

Then,

- $\tilde{f}(X)$ depends only on the law of X.
- $\tilde{f}$ is convex.
- If $\mathcal{G}$ is a σ-subalgebra of $\mathcal{F}$, then

$$\tilde{f}(\mathbb{E}[X|\mathcal{G}]) \leq \tilde{f}(X).$$

E-convex functions

Definition

A tracial W^*-function $f = (f^{\mathcal{A}})_{\mathcal{A}}$ is said to be *E-convex* if

1. Each $f^{\mathcal{A}}$ is convex and lower semi-continuous on $L^2(\mathcal{A})^m_{\mathrm{sa}}$.
2. If $\iota : \mathcal{A} \to \mathcal{B}$ is a trace-preserving $*$-homomorphism inclusion and $E = \iota^* : \mathcal{B} \to \mathcal{A}$ is the corresponding trace-preserving conditional expectation, then $f^{\mathcal{A}}(E[X]) \leq f^{\mathcal{B}}(X)$.

E-convex functions

Definition

A tracial W^*-function $f = (f^{\mathcal{A}})_{\mathcal{A}}$ is said to be E-convex if

1. Each $f^{\mathcal{A}}$ is convex and lower semi-continuous on $L^2(\mathcal{A})^m_{\text{sa}}$.
2. If $\iota : \mathcal{A} \to \mathcal{B}$ is a trace-preserving $*$-homomorphism inclusion and $E = \iota^* : \mathcal{B} \to \mathcal{A}$ is the corresponding trace-preserving conditional expectation, then $f^{\mathcal{A}}(E[X]) \leq f^{\mathcal{B}}(X)$.

Lemma

Let f be a E-convex tracial W^*-function. Then, for any $\mathcal{A}$ and $X \in L^2(\mathcal{A})^m_{\text{sa}}$, there exists a subgradient $Y \in L^2(\mathrm{W}^*(X))^m_{\text{sa}}$ of $f^{\mathcal{A}}$ at X.

Non-commutative duality

For a non-commutative law μ and tracial W^*-function f, define

$$\mu(f) := f^{\mathcal{A}}(X),$$

where $(\mathcal{A}, X)$ is a GNS construction of μ.

Non-commutative duality

For a non-commutative law μ and tracial W^*-function f, define

$$\mu(f) := f^{\mathcal{A}}(X),$$

where $(\mathcal{A}, X)$ is a GNS construction of μ.

Proposition (GJNS)

We have the following duality:

$$\sup_{(X,Y)\text{ coupling}} \langle X, Y\rangle_{L^2(\mathcal{A})^m_{\mathrm{sa}}} = \inf \big(\mu(f) + \nu(g)\big),$$

where inf is taken over admissible pairs (f, g) of E-convex functions:

$$f^{\mathcal{A}}(X_1) + g^{\mathcal{A}}(X_2) \geq \langle X_1, X_2\rangle_{L^2(\mathcal{A})^m_{\mathrm{sa}}}, \qquad \forall X_1, X_2 \in L^2(\mathcal{A})^m_{\mathrm{sa}}.$$

Non-commutative duality

Proposition (GJNS)

Let $(\mathcal{A}, X, Y)$ be a coupling of μ and ν. The following are equivalent.

1. $(\mathcal{A}, X, Y)$ is an optimal coupling.
2. There exists an admissible pair (f, g) such that
$$f^{\mathcal{A}}(X) + g^{\mathcal{A}}(Y) = \langle X, Y \rangle_{L^2(\mathcal{A})^m_{\mathrm{sa}}}.$$
3. There exists an E-convex W^*-function f such that $Y \in \eth f^{\mathcal{A}}(X)$.

Inf-convolution

- Inf-convolution: for tracial W^*-functions f and g,

$$(f \square g)^{\mathcal{A}}(X) := \inf_{\iota:\mathcal{A}\to\mathcal{B}, Y \in L^2(\mathcal{B})_{\mathrm{sa}}^m} \left\{ f^{\mathcal{B}}(\iota(X) - Y) + g^{\mathcal{B}}(Y) \right\}.$$

$f \square g$ is a tracial W^*-function.

Inf-convolution

- Inf-convolution: for tracial W^*-functions f and g,

$$(f \square g)^{\mathcal{A}}(X) := \inf_{\iota:\mathcal{A}\to\mathcal{B}, Y \in L^2(\mathcal{B})^m_{\mathrm{sa}}} \left\{ f^{\mathcal{B}}(\iota(X) - Y) + g^{\mathcal{B}}(Y) \right\}.$$

  $f \square g$ is a tracial W^*-function.
- Example: $q_t^{\mathcal{A}}(X) = (1/2t)\|X\|^2_{L^2(\mathcal{A})^m_{\mathrm{sa}}}$. It satisfies $q_s \square q_t = q_{s+t}$ and

$$q_s \square (q_t \square f) = (q_s \square q_t) \square f = q_{s+t} \square f.$$

  $(q_t \square f)_{t>0}$ is a W^*-analog of Hopf-Lax semigroup:

$$\begin{cases} u_t + \frac{1}{2}\|\nabla u\|^2 = 0, \\ u(0, \cdot) = f. \end{cases}$$

Regularization of Hopf-Lax semigroup

Let $q_t^{\mathcal{A}}(X) = (1/2t)\|X\|^2_{L^2(\mathcal{A})^m_{\mathrm{sa}}}$ for $t > 0$.

Proposition (GJNS)

Let f be an E-convex function. Then, the following are equivalent.

1. $f = q_t \square g$ for some E-convex g.
2. $q_t - f$ is E-convex.
3. $\mathcal{L}f - q_{1/t}$ is E-convex.
4. $\eth f(X)$ consists of a single point $\nabla f(X) \in L^2(\mathrm{W}^*(X))^m_{\mathrm{sa}}$, and ∇f defines a $1/t$-Lipschitz function.

$\mathcal{L}f$ denotes Legendre transform:

$$\mathcal{L}f^{\mathcal{A}}(X) = \sup_{\iota:\mathcal{A}\to\mathcal{B},\, Y\in L^2(\mathcal{B})^m_{\mathrm{sa}}} \{\langle \iota(X), Y\rangle - f^{\mathcal{B}}(Y)\}.$$

Applications — isomorphism of von Neumann algebras

Let $(\mathcal{A}, X, Y)$ be an optimal coupling of $\mu, \nu \in \Sigma_{m,R}$.
Then, $t \mapsto (1-t)X + tY$ is a geodesic.

Applications — isomorphism of von Neumann algebras

Let $(\mathcal{A}, X, Y)$ be an optimal coupling of $\mu, \nu \in \Sigma_{m,R}$.
Then, $t \mapsto (1-t)X + tY$ is a geodesic.

Theorem (GJNS)

We have $\mathrm{W}^*((1-t)X + tY) = \mathrm{W}^*(X, Y)$ for $t \in (0,1)$.

Applications — isomorphism of von Neumann algebras

Let $(\mathcal{A}, X, Y)$ be an optimal coupling of $\mu, \nu \in \Sigma_{m,R}$.
Then, $t \mapsto (1-t)X + tY$ is a geodesic.

Theorem (GJNS)

We have $\mathrm{W}^*((1-t)X + tY) = \mathrm{W}^*(X, Y)$ for $t \in (0,1)$.

(Proof) $Y \in \eth f(X)$ for some E-convex f. Let

$$X_t = (1-t)X + tY, \qquad f_t(X) = (1-t)\frac{1}{2}\|X\|_{L^2}^2 + tf(X).$$

Then,

$$X_t \in \eth f_t(X) \Longrightarrow X \in \eth \mathcal{L} f_t(X_t).$$

$q_{1-t} - \mathcal{L} f_t$ is E-convex, which implies $X \in L^2(\mathrm{W}^*(X_t))_{\mathrm{sa}}^m$.

Thank you.