Optimal Transport and Control

Ludovic Rifford

Université Nice Sophia Antipolis
&
Institut Universitaire de France

Workshop on Control and Observation of Nonlinear Control Systems with Application to Medicine

University of Hawaii at Manoa (Honolulu, September 5-7, 2013)
Let $M$ be a smooth manifold and $\mu_0$ and $\mu_1$ be probability measures on $M$. We call transport map from $\mu_0$ to $\mu_1$ any measurable map $T : M \rightarrow M$ such that $T_\#\mu_0 = \mu_1$, that is

$$\mu_1(B) = \mu_0(T^{-1}(B)), \quad \forall B \text{ measurable } \subset M.$$
The Monge Optimal Transport Problem

Let \( c : M \times M \to \mathbb{R} \) be a cost and \( \mu_0, \mu_1 \) two probability measures on \( M \), find a transport map \( T : M \to M \) from \( \mu_0 \) to \( \mu_1 \) which minimizes the transportation cost

\[
\int_M c(x, T(x)) d\mu_0(x).
\]

Existence ?

Uniqueness ?

Regularity ?
Example 1: Atomic measures

Let \( x, y_1, y_2 \in M \) with \( y_1 \neq y_2 \). Then there is no transport map from

\[
\mu_0 = \delta_x \quad \text{to} \quad \mu_1 = \frac{1}{2} \delta_{y_1} + \frac{1}{2} \delta_{y_2}.
\]
Example 2: The original Monge problem in $\mathbb{R}$

Given two probability measures $\mu_0, \mu_1$ in $\mathbb{R}$, we are concerned with transport maps $T : \mathbb{R} \to \mathbb{R}$ from $\mu_0$ to $\mu_1$ which minimize the transportation cost

$$\int_{\mathbb{R}} |T(x) - x| \, d\mu_0(x).$$

both optimal with the same transportation cost.

$T(x) = x + 1$ and $T(x) = 2 - x$
Example 3: The quadratic Monge problem in $\mathbb{R}^n$

Given two probability measures $\mu_0, \mu_1$ in $\mathbb{R}^n$, we are concerned with transport maps $T : \mathbb{R}^n \to \mathbb{R}^n$ from $\mu_0$ to $\mu_1$ which minimize the transportation cost

$$\int_{\mathbb{R}^n} |T(x) - x|^2 \, d\mu_0(x).$$

Theorem (Brenier ’91)

If $\mu_0$ is compactly supported and absolutely continuous with respect to the Lebesgue measure, there exists a unique optimal transport map with respect to the quadratic cost. In fact, there is a convex function $\psi : M \to \mathbb{R}$ such that

$$T(x) = \nabla \psi(x) \quad \mu_0 \text{ a.e. } x \in \mathbb{R}^n.$$
Example 4: The McCann Theorem

Let \((M, g)\) be a smooth compact Riemannian manifold equipped with its geodesic distance \(d_g\).
Example 4: The McCann Theorem

Let \((M, g)\) be a smooth compact Riemannian manifold equipped with its geodesic distance \(d_g\). Given two probability measures \(\mu_0, \mu_1\) on \(M\), we are concerned with transport maps \(T : M \to M\) which minimize the transportation cost

\[
\int_M d_g(x, T(x))^2 d\mu_0(x).
\]

**Theorem (McCann ’01)**

If \(\mu_0\) is absolutely continuous w.r.t. Lebesgue, there exists a unique optimal transport map from \(\mu_0\) to \(\mu_1\). In fact, there is a \(c\)-convex function \(\varphi : M \to \mathbb{R}\) satisfying

\[
T(x) = \exp_x (\nabla \varphi(x)) \quad \mu_0 \text{ a.e. } x \in \mathbb{R}^n.
\]
Example 5: A Monge problem with obstacle

Let $C$ be a smooth convex body in $\mathbb{R}^n$ and $d$ the geodesic distance on $\Omega := \mathbb{R}^n \setminus C$.

Given two probability measures $\mu_0, \mu_1$ in $\mathbb{R}^n$, we are concerned with transport maps $T : \mathbb{R}^n \to \mathbb{R}^n$ from $\mu_0$ to $\mu_1$ which minimize the transportation cost $\int_{\mathbb{R}^n} d(x, T(x))^2 d\mu_0(x)$.

Theorem (Cardaliaguet-Jimenez ’11)

Existence (but not uniqueness in general).
The purpose of this talk is to study optimal transport problems with **costs coming from optimal control problems**. Two types of costs:

- LQR costs
- Quadratic sub-Riemannian distances.
The Kantorovitch Optimal Transport Problem

Given $M$, a cost $c : M \times M \rightarrow \mathbb{R}$ and two probability measures $\mu_0, \mu_1$ on $M$, we want to find a probability measure $\gamma$ on $M \times M$ having marginals $\mu_0$ and $\mu_1$, i.e.

$$(\pi_1)_\# \gamma = \mu_0 \quad \text{and} \quad (\pi_2)_\# \gamma = \mu_1,$$

(where $\pi_1 : M \times M \rightarrow M$ and $\pi_2 : M \times M \rightarrow M$ are the canonical projections), which minimizes the transportation cost given by

$$\int_{M \times M} c(x, y) d\gamma(x, y).$$

When the transport condition $(\pi_1)_\# \gamma = \mu, (\pi_2)_\# \gamma = \nu$ is satisfied, we say that $\gamma$ is a transport plan, and if $\gamma$ minimizes also the cost we call it an optimal transport plan.
Kantorovitch allows splitting

$$\mu_0(B) = \gamma(B \times M)$$
Let $M$ be a smooth manifold, $c : M \times M \to \mathbb{R}$ be a cost function, and $\mu_0, \mu_1$ two probability measures on $M$.

**Monge’s Problem**

Minimize

$$\int_M c(x, T(x)) d\mu_0(x)$$

among all transport maps $T$, that is $T_\#\mu_0 = \mu_1$.

**Kantorovitch’s Problem**

Minimize

$$\int_M c(x, y) d\gamma(x, y)$$

among all transport plans $\gamma$, that is $(\pi_1)_\#\gamma = \mu_0, (\pi_2)_\#\gamma = \mu_1$. 

---

Ludovic Rifford

Optimal Transport and Control
Reminder: The Brenier-McCann Theorem

**Theorem (Brenier ’91)**

If $\mu_0$ is compactly supported and absolutely continuous with respect to the Lebesgue measure, there exists a unique optimal transport map with respect to the quadratic cost. In fact, there is a **convex** function $\psi : M \to \mathbb{R}$ such that

$$T(x) = \nabla \psi(x) \quad \mu_0 \text{ a.e. } x \in \mathbb{R}^n.$$ 

**Theorem (McCann ’01)**

If $\mu_0$ is absolutely continuous w.r.t. Lebesgue, there exists a unique optimal transport map from $\mu_0$ to $\mu_1$. In fact, there is a **c-convex** function $\varphi : M \to \mathbb{R}$ satisfying

$$T(x) = \exp_x (\nabla \varphi(x)) \quad \mu_0 \text{ a.e. } x \in \mathbb{R}^n.$$
Kantorovitch’s Duality

Theorem

There are two continuous functions \( \psi_1, \psi_2 : M \to \mathbb{R} \) satisfying

\[
\psi_1(x) = \max_{y \in M} \{\psi_2(y) - c(x, y)\} \quad \forall x \in M,
\]

\[
\psi_2(y) = \min_{x \in M} \{\psi_1(x) + c(x, y)\} \quad \forall y \in M.
\]

such that the following holds: a transport plan \( \gamma \) is optimal if and only if one has

\[
\psi_2(x) - \psi_1(y) = c(x, y) \quad \text{for } \gamma \text{ a.e. } (x, y) \in M \times M.
\]

As a consequence, to obtain that an optimal transport plan corresponds to a Monge’s optimal transport map, we have to show that \( \gamma \) is concentrated on a graph.
Proof of Brenier-McCann’s Theorem I

Returning to the Riemannian setting, let $\psi_1, \psi_2 : M \to \mathbb{R}$ be a pair of Kantorovitch potentials given by the previous result.

- The function $x \mapsto d_g(x, y)^2$ is locally Lipschitz on $M$.
- The function $\psi_1$ is locally Lipschitz on $M$. As a consequence, by Rademacher’s Theorem, it is differentiable $\mu_0$-a.e.
- Let $\bar{x} \in \text{supp}(\mu_0)$ be such that $\psi_1$ is differentiable at $\bar{x}$. Let $\bar{y}$ be such that

$$\psi_1(\bar{x}) = \psi_2(\bar{y}) - d_g(\bar{x}, \bar{y})^2.$$ 

Then we have,

$$d_g(x, \bar{y})^2 \geq \psi_2(\bar{y}) - \psi_1(x) \quad \forall x \in M.$$
Proof of Brenier-McCann’s Theorem II

Any Lipschitz curve $c : [0, 1] \rightarrow M$ with $c(1) = \bar{y}$ satisfies

$$
\int_0^1 |\dot{c}(t)|_{c(t)}^2 \, dt \geq d_g (c(0), \bar{y})^2 \geq \psi_2(\bar{y}) - \psi_1(c(0)),
$$

with equality if $c$ is the minimizing geodesic from $\bar{x}$ to $\bar{y}$. 

$$
\Rightarrow \quad \bar{y} = \exp_{\bar{x}} (\nabla \psi_1(\bar{x})).
$$
TWO ISSUES

(Regularity) Show that $\psi_1$ is differentiable $\mu$-a.e.

(Twist) Deduce that, if $\psi_1$ is differentiable at $\bar{x}$, then there is a unique $\bar{y}$ such that

$$\psi_1(\bar{x}) = \psi_2(\bar{y}) - c(\bar{x}, \bar{y}).$$
Let us consider the linear control system

\[ \dot{x} = Ax + Bu \]

where \( x \in \mathbb{R}^n \), \( u \in \mathbb{R}^m \) and \( A, B \) are \( n \times n \) and \( n \times m \) matrices. Cost

\[ c(x, y) = \inf \left\{ \int_0^1 L(x(t), u(t)) \, dt \mid u \in L^2, x_u(0) = x, x_u(1) = y \right\} \]

with \((W\) sym. nonneg. and \( U\) sym. def. pos)

\[ L(x, u) = \langle x, Wx \rangle + \langle u, Uu \rangle \]

**Theorem (Hindawi-Pomet-R ’11)**

Existence, uniqueness, and regularity.
Let \((M, \Delta, g)\) be a complete sub-Riemannian structure of dimension \(n\) and rank \(m < n\), that is

- \(M\) a smooth connected manifold.
- \(\Delta\) a totally nonholonomic distribution.
- \(g\) a smooth metric on \(\Delta\).

Let \(d_{SR}(\cdot, \cdot)\) be the sub-Riemannian distance on \(M \times M\), i.e.

\[
d_{SR}(x, y) = \inf \left\{ \text{length}_g(\gamma) \mid \gamma \in \mathcal{C}^1, \gamma(0) = x, \gamma(1) = y, \dot{\gamma}(t) \in \Delta_{\gamma(t)} \right\}.
\]

From now on, we assume that the metric space \((M, d_{SR})\) is complete.
Let $\mu, \nu$ be two compactly supported probability measures on $M$. Find a measurable map $T : M \to M$ satisfying

$$T_\# \mu = \nu,$$

and in such a way that $T$ minimizes the transportation cost given by

$$\int_M d_{SR}(x, T(x))^2 d\mu(x).$$

**Theorem (Figalli-R ’10)**

Assume that there exists an open set $\Omega \subset M \times M$ such that $\text{supp}(\mu \times \nu) \subset \Omega$, and $d_{SR}^2$ is locally Lipschitz on $\Omega \setminus D$. Then, there is existence and uniqueness of an optimal transport map.
Examples

- Example 1: Two generating distributions

  Proposition (A. Agrachev, P. Lee, 2008)

  If $\Delta$ is two-generating on $M$, then the squared sub-Riemannian distance function is locally Lipschitz on $M \times M$.

- Example 2: Generic sub-Riemannian structures

  Proposition (Y. Chitour, F. Jean, E. Trélat, 2006)

  Let $(M, g)$ be a complete Riemannian manifold of $\text{dim} \geq 4$. Then, for any generic distribution of rank $\geq 3$, the squared sub-Riemannian distance function is locally semiconcave (hence locally Lipschitz) on $M \times M \setminus D$. 
Example 3: Rank-two distributions in dimension 4
Consider the distribution $\Delta$ in $\mathbb{R}^4$ spanned by

\begin{align*}
f_1 &= \partial_{x_1}, \\
f_2 &= \partial_{x_2} + x_1 \partial_{x_3} + x_3 \partial_{x_4}.
\end{align*}

A horizontal path $\gamma : [0, 1] \to \mathbb{R}^4$ is singular if and only if it satisfies (up to reparameterization by arc-length)

\[ \dot{\gamma}(t) = f_1(\gamma(t)), \quad \forall t \in [0, 1]. \]

Then, for any metric, the sub-Riemannian distance function $d_{SR}$ is locally lipschitz on the set

\[ \Omega = \{(x, y) \in \mathbb{R}^4 \times \mathbb{R}^4 \mid (y - x) \notin \text{SPAN}\{e_1\} \} \]

Consequently, we get existence and uniqueness of optimal transport maps.
Thank you for your attention!!