Report for CSE 5339 2018 — (OTMLSA)
Optimal Transport in Machine Learning and Shape Analysis

BRENIER'S POLAR FACTORIZATION THEOREM AND MCCANN'S GENERALIZATION

OSMAN BERAT OKUTAN

4/25/2018

1 Introduction

Brenier's polar factorization theorem is a factorization theorem for vector valued functions on Euclidean domains, which generalizes classical factorization results like polar factorization of real matrices and Helmotz decomposition of vector fields.

Theorem 1.1 (Brenier's polar factorization theorem)**.** *[1] Given a probability space (X, μ) and a bounded domain Ω in $\mathbb{R}^d$ equipped with the Lebesgue measure $|\cdot|$ (normalized so that $|\Omega| = 1$) [1] for every vector-valued function $u \in L_p(X, \mu; \mathbb{R}^d)$ there is a unique "polar factorization" $u = \nabla\psi \circ s$, where ψ is a convex function defined on Ω and s is a measure-preserving mapping from (X, μ) into $(\Omega, |\cdot|)$, provided that u is nondegenerate, in the sense that $\mu(u^{-1}(E)) = 0$ for each Lebesgue negligible subset E of $\mathbb{R}^d$.*

The proof is obtained by using a proper Monge-Kantorovich problem.

McCann generalized Brenier's factorization theorem to functions on Riemannian manifolds. The statement is as follows:

Theorem 1.2 ([2])**.** *Let M be a connected compact Riemannian manifold. Let $s : M \to M$ be a Borel map which never maps positive volume into zero volume. Then s factors uniquely into the form $s = t \circ u$, where*

$$u : M \to M \text{ is a volume preserving map and}$$

$$t = \exp(\nabla\psi) : M \to M$$

where ψ is a "convex" function $\psi : M \to \mathbb{R}$.

Here, the definition of convexity is more technical. Let $c : M \times M \to \mathbb{R}$ be a function. A function $\psi : M \to \mathbb{R}$ is called c-concave if $\psi(y) = \inf_{x \in M} c(x, y) - \psi(x)$ for each $y \in M$. A c-convex function is a real valued function on M whose negative is c-concave. For the statement above, we take $c = d^2(x, y)/2$ where d is the Riemannian distance. The proof of Theorem 1.2 is still obtained through a Monge-Kantorovich problem, but now in a more general setting.

In the following sections, we review Monge and Kantorovich problems and then give the idea of the proof of the factorization theorem.

[1] under additional technical assumptions on X and Ω

2 Monge problem

Let $(X, \mu), (Y, \nu)$ be measure spaces. Let $c : X \times Y \to \mathbb{R}_{\geqslant 0}$ be a function. In the setting of optimal transportation, we interpret these as follows:

- X as the collection of distributors where the amount of material that each distributor has is determined by μ. More precisely, the distributor $x \in X$ has $d\mu(x)$ material.
- Y as the collection of receivers where the amount each receiver needs is determined by ν, i.e. the receiver y needs $d\mu(y)$ material.
- $c(x, y)$ is the cost of transporting one unit material from x to y.

Definition 1. *A transportation map from X to Y is a measurable map $G : X \to Y$ such that the pusforward measure $G_*(\mu) = \nu$. In other words for each measurable set $U \in Y$, $\nu(U) = \mu(G^{-1}(U))$.*

The interperation of a transportation map is follows: if $G(x) = y$ then all materials of the distributer x should be sent to the recevier y. The pusforward condition guarantees that each $y \in Y$ receives exactly the amount it needs (note that pushforward along G can be interpreted as summation/integration over each fiber of G). The total cost of such transportation should be the sum of individual costs, where the cost of sending all material of the distributor x to $G(x)$ is $c(x, G(x))d\mu(x)$. Therefore, we give the following definition:

Definition 2. *The total cost of a transportation plan $G : X \to Y$ is defined by*

$$C(G) := \int_X c(x, G(x))d\mu(x).$$

The Monge problem is finding the cost minimizing transportation plan. Such a plan is called a Monge solution. Existence and uniqueness of a Monge solution depends on the properties of the individual cost function $c : X \times Y \to \mathbb{R}$.

3 Kantorovich problem

Note that a transport plan does not allow a distributor to distribute its material among different receivers. In the Kantorovich setting this is allows. Note that in this case, instead of a map $G : X \to Y$, one needs a mathematical object γ such that $\gamma(x, y)$ signifies the amount transported from x to y, and this distribution should be such that summing over x should give $d\nu(y)$ (hence y gets what it needs) and summing over y should give $d\mu(x)$ (hence x distributes everything it has). The mathematical object fitting to this description is called a coupling.

Definition 3. *A coupling between $(X, \mu), (Y, \nu)$ is a measure γ on X, Y whose pushforward to X (resp. Y) under the canonical projection map is μ (resp. ν).*

Definition 4. *A transportation plan from X to Y is a coupling γ between $(X, \mu), (Y, \nu)$.*

Note that in this setting, $d\gamma(x, y)$ denotes the amount of material transported from x to y. The cost of this individual transportation is $c(x, y)d\gamma(x, y)$. Hence, for the total cost of a transportation plan we give the following definition:

Definition 5. *The total cost of the transportation plan γ is defined by*

$$C(\gamma) := \int_{X \times Y} c(x, y)d\gamma(x, y).$$

The Kantorovich problem is finding the cost minimizing transportation plan. Such a transportation plan is called a Kantorovich solution.

Remark 3.1. *The Kantorovich problem is a relaxation of the Monge problem in the following sense: Given a transportation plan $G : X \to Y$, the measure $\gamma := (id_X \times G)_*(\mu)$ is a transportation plan such that $C(G) = C(\gamma)$. Furthermore, the support of γ is the graph of G.*

Remark 3.2. *If μ, ν are measures on the Borel sigma algebra, then the collection of transportation can be considered as a convex subset of a Banach space, namely the dual space of the space of real valued continuous functions on $X \times Y$ with l_∞-norm. Hence, the possible candidates for the Kantorovich problem, unlike the candidates for the Monge problem, has a linear structure which helps in showing existence and uniqueness of solutions. Furthermore, the functional to be minimized is linear in its inputs for the Kantorovich problem.*

A very important aspect of the Kantorovich problem is the following duality (see [3, Section 2.1]:

$$\min_{\gamma \text{ coupling between } \mu,\nu} \int c d\gamma = \sup_{(-u,-v)\in Lip_c} -\int_X u d\mu - \int_Y v d\nu,$$

where

$$Lip_c = \{(u, v) : u, v \text{ are } L^1, c(x, y) \geqslant u(x) + v(y) \forall (x, y) \in X \times Y\}.$$

This is called the Kantarovich duality and crucial in obtaining Monge-Kantarovich solutions.

4 Idea of proof of Theorem 1.2

In this section we assume that X is a compact connected Riemannian manifold, $Y = X$ and $c(x, y) = d^2(x, y)/2$ where d denotes the Riemannian distance. This cost function, compared to the cost function given by the distance, is better suited for obtaining a Monge-Kantarovich solution because of its proper *differentiability* properties (see [2, Proposition 6]).

Note that when $X = Y$ are compact connected Riemannian manifolds, a transportation plan can also be described by a vector field as follows: Let G be a transportation plan $G : X \to Y$. For each $x \in X$ choose a unit speed geodesic α_x from x to $G(x)$. Let V be the vector field on X defined by $V(x) = d(x, G(x))\alpha_x'(0)$. Hence, $G(x) = \exp(V(x))$. Therefore $G = \exp(V)$. A question one can ask is the following: Does such a V arise as the gradient of a potential function? The following theorem answers this question positively and it is the main result about Monge-Kantorovich problem used in the proof of Theorem 1.2.

Theorem 4.1 (Existence of Monge solutions, uniqueness of Kantorovich solutions)**.** *[3, Theorem 2.9] Let M be an n-dimensional connected compact Riemannian manifold, and μ, ν be Borel measures on M. Then there is a convex potential function $\psi : M \to \mathbb{R}$ such that*

1. *G(x):= $\exp_x(\nabla\psi)$ is a transport map.*
2. *G is the only transport map arising this way. It solves the Monge problem.*
3. *The Kantorovich problem has a unique solution.*
4. *The Kantorovich solution is attained through G.*

Hence in this case, the Monge problem and the Kantorovich problem has the same unique solution.

Now we can start discussing the sketch of the proof of Theorem 1.2.

Idea of proof of Theorem 1.2. Let μ be the Riemanniam volume measure on M and let $\nu = s_*(\mu)$. By Theorem 4.1, there exist a solution t of the Monge problem from (X, μ) to (X, ν) which is the exponential of the gradient of a convex potential function $\psi : M \to \mathbb{R}$.

Let t^* be the solution of the Monge problem from (X, ν) to (X, μ), whose existence is guaranteed by Theorem 4.1. Show that t, t^* are inverses almost everywhere. Let $u = t^* \circ s$. Then $t \circ u = s$, μ almost everywhere. Furthermore $u_*(\mu) = t^*_* s_*(\mu) = t^*_*(\nu) = \mu$, hence u is measure preserving. $\square$

References

[1] Brenier, Yann. "Polar factorization and monotone rearrangement of vector-valued functions." Communications on pure and applied mathematics 44.4 (1991): 375-417.

[2] McCann, Robert J. "Polar factorization of maps on Riemannian manifolds." Geometric & Functional Analysis GAFA 11.3 (2001): 589-608.

[3] McCann, Robert J., and Nestor Guillen. "Five lectures on optimal transportation: geometry, regularity and applications." Analysis and geometry of metric measure spaces: lecture notes of the séminaire de Mathématiques Supérieure (SMS) Montréal (2011): 145-180.