ON THE MONGE PROBLEM

The first thing to point out about the MONGE problem is that, given arbitrary probability measures μ on $(X, \mathcal{F})$ and ν on $(Y, \mathcal{G})$, a map $T: X \to Y$ that transports μ to ν may not exist.

Indeed, the transport property $T \# \mu = \nu$ is equivalent to

(1) $$\int_Y \varphi(y)\, \nu(dy) = \int_X \varphi(T(x))\, \mu(dx), \quad \forall\ \varphi \text{ bdd, meas.}$$

So, if $\mu = \delta_{x_0}$ is a DIRAC measure for some x_0, this becomes $\int_Y \varphi(y)\, \nu(dy) = \varphi(T(x_0))$ and forces $\nu = \delta_{T(x_0)}$. In other words, unless the given ν is precisely this DIRAC measure on $(Y, \mathcal{G})$, a transport map does not exist.

THE REAL LINE: Let us concentrate for a moment on the case $X = Y = \mathbb{R}$ and on two probability measures μ, ν on $\mathcal{B}(\mathbb{R})$ with respective probability distribution functions

$$F(\cdot) \triangleq \mu((0, \cdot]), \quad G(\cdot) \triangleq \nu((0, \cdot])$$

and right-continuous inverses $F^{-1}(\cdot)$, $G^{-1}(\cdot)$, respectively.

When $\mu = \lambda|_{[0,1]}$ is LEBESGUE measure (uniform distribution) on the unit interval, then the map $T(x) = G^{-1}(x)$, $0 \le x \le 1$ transports μ to ν. This is the familiar SKOROKHOD construction.

. A bit more generally, <u>when μ has no atoms</u> (i.e., $F(\cdot)$ is continuous), the mapping

(2) $$T(x) \triangleq G^{-1}(F(x)), \quad x \in \mathbb{R}$$

transports μ to ν: i.e., $T \# \mu = \nu$. This can be verified easily.

The mapping in (2) above is called <u>monotone re-arrangement.</u>

We shall extend now this idea.

<u>SEVERAL DIMENSIONS</u>: Let us start with the Euclidean plane $X = Y = \mathbb{R}^2$

<u>THEOREM</u>: <u>DISINTEGRATION OF MEASURE</u>.

Fix $\mu \in \mathcal{P}(\mathbb{R}^2)$ and define $\mu_1 \triangleq P_1 \# \mu$, the "1-marginal" of μ, where

$$P_1(x_1, x_2) \triangleq x_1, \quad P_2(x_1, x_2) = x_2.$$

There exists then a family of probability measures $(\mu_{x_1})_{x_1 \in \mathbb{R}}$ on the real line, with

(3) $$\mu(dx_1, dx_2) = \mu_{x_1}(dx_2) \otimes \mu_1(dx_1).$$

Illustration: Conditional Distributions.

Suppose $\mu(A) = \iint_A f(x_1,x_2)\,dx_1\,dx_2$, $A \in \mathcal{B}(\mathbb{R}^2)$ for some $f > 0$ with $\iint_{\mathbb{R}^2} f = 1$. Define the "marginal density"

$$f_1(x_1) \triangleq \int_{\mathbb{R}} f(x_1,x_2)\,dx_2\,, \quad x_1 \in \mathbb{R}.$$

Then μ_1 is the "marginal probability measure" $\mu_1(\cdot) = (P_1 \# \mu)(\cdot) = \int_{\cdot} f_1(x_1)\,dx_1$

and

$$\mu_{x_1}(\cdot) = \int_{\cdot} \frac{f(x_1,x_2)}{f_1(x_1)}\,dx_2$$

the corresponding "conditional distribution".

Now, consider $\underline{\text{two}}$ such absolutely continuous measures on $\mathcal{B}(\mathbb{R}^2)$:

- $$\mu(dx_1,dx_2) = f(x_1,x_2)\,dx_1\,dx_2 = \underbrace{\frac{f(x_1,x_2)}{f_1(x_1)}\,dx_2}_{\mu_{x_1}(dx_2)} \otimes \underbrace{f_1(x_1)\,dx_1}_{\mu_1(dx_1)} \qquad (4)$$

- $$\nu(dy_1,dy_2) = g(y_1,y_2)\,dy_1\,dy_2 = \underbrace{\frac{g(y_1,y_2)}{g_1(y_1)}\,dy_2}_{\nu_{y_1}(dy_2)} \otimes \underbrace{g_1(y_1)\,dy_1}_{\nu_1(dy_1)} \qquad (5)$$

From our discussion about monotone re-arrangements, consider the mapping $T_1 : \mathbb{R} \to \mathbb{R}$ that transports μ_1 to

(6) ν_1, i.e., $T_1 \# \mu_1 = \nu_1$; and, for μ_1- a.e. $x_1 \in \mathbb{R}$, the monotone re-arangement $T_2(x_1, \cdot) : \mathbb{R} \to \mathbb{R}$ that transports

(7) μ_{x_1} to ν_{x_1} : $T_2(x_1, \cdot) \# \mu_{x_1} = \nu_{x_1}$.

<u>THEOREM</u> (KNOTHE, 1957): The map

(8) $$\mathbb{R}^2 \ni (x_1, x_2) \longmapsto T(x_1, x_2) = \big(T_1(x_1), T_2(x_1, x_2) \big) \in \mathbb{R}^2$$

transports μ to ν.

Now iterate this procedure, to extend to $\mathbb{R}^3$, ... et cetera.

Here is BRENIER's theorem, the first general result about the solvability of the MONGE problem.

<u>THEOREM</u> (BRENIER, 1987): Let $X = Y = \mathbb{R}^n$ and consider two probability measures μ, ν with

$$\int_{\mathbb{R}^n} |x|^2 \mu(dx) + \int_{\mathbb{R}^n} |y|^2 \nu(dy) < \infty .$$

There exists then a (unique) optimal coupling $\bar{\pi} \in \Pi(\mu, \nu)$ for the KANTOROVICH problem with $c(x,y) = |x-y|^2$, and this is of the form

$$\text{(9)} \qquad \bar{\pi} = \left(Id \times T_* \right) \# \mu \; , \quad T_* \triangleq \nabla F$$

for some convex $F : \mathbb{R}^n \to \mathbb{R}$. In particular, the map T_* in (9) is then optimal for the corresponding MONGE problem, whose value is then the same as that of the KANTOROVICH problem:

$$\underbrace{I(\bar{\pi}) = \mathcal{T}_c(\mu,\nu)}_{\text{optimality of } \bar{\pi} \text{ for KANTOROVICH}} \underbrace{=}_{\text{equality of the two values}} \mathcal{M}_c(\mu,\nu) \underbrace{=}_{\text{optimality of } T_* \text{ for MONGE}} G(T_*) = \int_{\mathbb{R}^n} |x - \nabla F(x)|^2 \mu(dx).$$

MONGE – AMPÈRE EQUATION

Consider the setting of BRENIER's Theorem (first sentence) with the added feature that both measures are absolutely continuous w.r.t. LEBESGUE with densities $f > 0$, $g > 0$:

$$\mu(\cdot) = \int_{\cdot} f(x)\, dx \quad , \quad \nu(\cdot) = \int_{\cdot} g(x)\, dx .$$

We know from that result that μ can be transported to ν via the gradient $T = \nabla F$ of a convex function $F : \mathbb{R}^n \to \mathbb{R}$, i.e., $T \# \mu = \nu$; equivalently,

(10) $$\int_{\mathbb{R}^n} \varphi(y) \cdot \underbrace{g(y)\, dy}_{\nu(dy)} = \int_{\mathbb{R}^n} \varphi(\nabla F(x)) \cdot \underbrace{f(x)\, dx}_{\mu(dx)}$$

for every $\varphi \in C_b(\mathbb{R}^n)$.

And let us assume, for concreteness, that the BRENIER function F is smooth, say of class $C^1(\mathbb{R}^n)$, and strictly convex; then we can write the LHS of (10), via the change-of-variable $y = \nabla F(x)$, as

$$\int_{\mathbb{R}^n} \varphi(y)\, g(y)\, dy = \int_{\mathbb{R}^n} \varphi(\nabla F(x))\, g(\nabla F(x)) \det\left(D^2 F(x)\right) \cdot dx .$$

Comparing with (10), and because φ is arbitrary, we get

(11) $$\det\left(D^2 F(x)\right) = \frac{f(x)}{g(\nabla F(x))} \quad , \quad x \in \mathbb{R}^d ,$$

a partial differential (non-linear) equation for the BRENIER map F. It is a special case of the famous MONGE – AMPÈRE equation

$$\det\left(D^2 F(x)\right) = H\left(x, F(x), \nabla F(x)\right), \quad x \in \mathbb{R}^n.$$