Lecture 6

Polar Factorization

RECALL

- optimal map looks like gradient of convex func.

Brenier's Polar Factorization Theorem

trying to factor some kind of vector valued mapping

- given a reasonable vector field we can factor it as the composition of the gradient of a convex function & some other measure-preserving rearrangement

Let $\Omega \subset \mathbb{R}^n$ be open & bounded, w/ $|\partial\Omega| = 0$ (a non ridiculous boundary). Let

$r : \Omega \to \mathbb{R}^n$ be any non degenerate, L^2 vector valued mapping

- i.e if something has zero measure then it originated from a set of zero measure

$$|N| = 0 \implies |r^{-1}(N)| = 0$$

Then, thm says $\exists$ a unique convex function $\Psi : \Omega \to \mathbb{R}$ (up to additive constants) and a measure-preserving rearrangement

$S : \Omega \to \Omega$ s.th

$$r(x) = \nabla \Psi(S(x))$$

if r happens to be an optimal transport map then we can take S to be the identity map & we ~~need~~ knew that an optimal transport map can be written as the gradient of a convex func.

IDEA: Take Ω & rearrange & relabel so that @ the end of the day r is effectively the gradient of a convex function

PROOF

- we want to try to modify a transformed dual problem $(DP)^*$ to study this

minimize $\hat{I}[\varphi, \psi] = \int_\Omega [\varphi(r(x)) + \psi(x)]\, dx$

subject to $\varphi(x) + \psi(y) \geq x \cdot y \quad \forall x, y$

PROOF (continued)

· assuming φ & ψ are a convex dual pair, as before

- is $\hat{L}$ bounded below

$$\hat{L}[\varphi,\psi] \geq \int_\Omega r(x)\cdot x \, dx$$ ← in order to place bounds on this sort of object need Cauchy-Schwarz

$$\geq -\sqrt{\int_\Omega -|r(x)|^2 dx \int_\Omega |x^2| \, dx}$$ ← this is finite since $r \in L^2$ & Ω is bounded

∴ There is a finite infimum

· Now, we ask is the infimum attained (i.e. a min)

~~There~~ Like before, there is also a minimum by Arzelà-Ascoli

· want to extract some measure preserving map, S, from this problem

⇒ want to find $S(x)$ s.th

$$\int_\Omega h(x)\,dx = \int_\Omega h(S(x))\,dx \quad \forall\, h \in C^0(\bar{\Omega})$$

- Let's perturb φ, ψ as before:

φ

$$\psi_\varepsilon(y) = \psi(y) + \varepsilon h(y)$$

$$\varphi_\varepsilon(x) = \sup_y \{x\cdot y - \psi_\varepsilon(y)\}$$

This is an admissible pair b/c $\varphi_\varepsilon(x)$ is a Legendre transform of $\psi_\varepsilon(y)$

- Now we've perturbed our optimal pair, so it probably is not optimal any more continued...

$$\frac{\hat{L}[\varphi_\varepsilon, \psi_\varepsilon] - \hat{L}[\varphi, \psi]}{\varepsilon} \geq 0$$ this is the min

$$= \int_\Omega \frac{\varphi_\varepsilon(v(x)) - \varphi(v(x))}{\varepsilon} dx + \int_\Omega h(y)\, dy$$

need to make argument about what this will do in the limit

NOTE $x \cdot y - \psi(y)$ is maximized when $y = \nabla\varphi(x)$

$$\varphi(v(x)) = \sup_y \{v(x) \cdot y - \psi(y)\}$$ this is max when $y = \nabla\varphi(x)$

$$= v(x) \cdot \nabla\varphi(v(x)) - \psi[\nabla\varphi(v(x))]$$

now need to estimate $\varphi_\varepsilon(v(x)) \Rightarrow$ need to max $\underbrace{x \cdot y - \psi(y) - \varepsilon h(y)}_{\text{max when } y = \nabla\varphi(x) + o(1)}$

"something small"

$$\varphi_\varepsilon(v(x)) = v(x) \cdot (\nabla\varphi(v(x)) + o(1))$$
$$- \psi(\nabla\varphi(v(x)) + o(1)) - \varepsilon h(\nabla\varphi(v(x)) + o(1))$$

As before, we argue that:

$$\lim_{\varepsilon \to 0} \int \frac{\varphi_\varepsilon(v(x)) - \varphi(v(x))}{\varepsilon} dx = -\int_\Omega h(\nabla\varphi(v(x)))\, dx$$

Do the same thing where h is replaced w/ $-h$

$$\therefore \int_\Omega h(\underbrace{\nabla\varphi(v(x))}_{S(x)})\, dx = \int_\Omega h(y)\, dy$$

we were looking to pull, out of this problem, a measure-preserving map

$\boxed{\therefore S(x) = \nabla\varphi(v(x)) \text{ is in fact measure-preserving}}$

continued...

φ & ψ are legendre duals, so there is symmetry in the way we can represent them

$\Rightarrow \boxed{v(x) = \nabla\psi(s(x))}$

Ex) Special case of vector fields ~~that~~ close to the identity (Going to do asymptotics)

Let m be any nice enough vector field and let $v(x) = x + \varepsilon m(x)$

$$v(x) = x + \varepsilon m(x)$$

$$= \nabla\psi(s(x))$$

for some convex ψ & measure-preserving S.

How would we factor or decompose the identity map?

use $S_0(x) = x$, $\quad \psi_0(x) = |x^2|/2$

simplest measure-preserving map (Identity)

doesn't move anything

simplest convex func. (quadratic)

$$\nabla\psi_0(s_0(x)) = x$$

make the ansatz:

$$S(x) = x + \varepsilon S_1(x) + \underbrace{O(\varepsilon^2)}_{\text{higher order terms}}$$

$$\psi(x) = \frac{|x|^2}{2} + \varepsilon\,\psi_1(x) + O(\varepsilon^2)$$

From polar factorization of v:

$$x + \varepsilon m(x) = (x + \varepsilon S_1(x) + O(\varepsilon^2)) + \varepsilon\nabla\psi_1(x + \varepsilon S_1(x) + O(\varepsilon^2)) + O(\varepsilon^2)$$

$$= x + \varepsilon\left[S_1(x) + \nabla\psi_1(S_1(x))\right] + O(\varepsilon^2)$$

$\Rightarrow \boxed{m(x) = S_1(x) + \nabla\psi_1(x)}$

We also know that S is measure-preserving

- from change of vars we know

$$1 = \det(\nabla S(x))$$

$$= \det\left(I + \varepsilon \nabla S_1(x) + O(\varepsilon^2)\right)$$

$$= 1 + \varepsilon \underbrace{\nabla \cdot S_1(x)}_{\text{divergence of } S_1} + O(\varepsilon^2)$$

divergence of S_1 is the sum of diag elements in det

$$\Rightarrow \boxed{\nabla \cdot S_1(x) = 0}$$

* We found that we can decompose ~~an arbr~~ a vector field, m, into the sum of a gradient & a divergence free field *

· helmholtz decomposition is a special case (linearization) of polar factorization

· Can we relate ~~this~~ polar factorization to an optimal transport. like minimization problem by working backwards?

we solved

$$\min_{\phi} \int \left(\phi(v(x)) + \psi(x)\right) dx \qquad \text{subject to } \phi(x) + \psi(y) \geq x \cdot y$$

integrating against a Lesbegue measure

~~can we make this look~~

· Let's make this look more like (DP)*

· We have freedom to change measures

* we want $\phi(*)$ integrated against a measure μ & $\psi(\,)$ integrated against a measure ν

→ Let ν be the Lebesgue measure

→ want a measure μ s.th $\int h(v(x))\,dx = \int h(y)\,d\mu(y) \quad \forall h$

* Let $\mu = v_{\#}\nu$ (push forward of Lebesgue through v) *

In this notation, we are studying:

$$\int \varphi(y)\, d\mu(y) + \int \psi(x)\, d\lambda(x)$$

subject to $\varphi(x) + \psi(y) \geq x \cdot y$

← This is EXACTLY our transformed dual prob $(DP)^*$ w/ appropriate measure μ & λ

⟹ There's an optimal transport problem behind this

Equivalent to quadratic cost OT b/w μ & λ (where λ is the Lebesgue measure)

We are solving:

$$\min \int |x - T(x)|^2 \, d\mu(x)$$

~~NOTE~~

$$\text{s.t. } T_{\#}\mu = \lambda$$

Where did mapping $s(x)$ go? And the $\nabla\psi$?

vector field that maps Ω to whatever

$$(\Omega, \lambda) \xrightarrow{\ r\ } (r(\Omega), \mu)$$

starting point

← push forward ~~measure~~

Then we did polar factorization.

polar factorization

$$\begin{tikzcd}
(\Omega, \lambda) \arrow[rr, "r"] \arrow[dr, "s"'] & & (r(\Omega), \mu) \arrow[dl, bend left, "T = \nabla\varphi(\,)"] \\
& (\Omega, \lambda) \arrow[ur, "\Delta\psi"] &
\end{tikzcd}$$

NOTE: S is just a measure-preserving rearrangement of Ω

$$r(x) = \nabla \psi(s(x))$$

$$s(x) = \nabla \varphi(r(x))$$

Let's go back to obj. func. & try to rearrange to see if polar factorization fits here

Object func. for OT is

$$\int |x - T(x)|^2 \, d\mu(x)$$

$$= \int |v(x) - T(v(x))|^2 dx$$

(def'n of μ)

- Again the freedom is in the measure
  - we also know that T is connected to the measure
- using the fact that

$$\int h(v(x))\,dx = \int h(y)\,d\mu(y)$$

$$= \int |v(x) - s(x)|^2 \, dx$$

b/c $T = \nabla\varphi$ & $s(x) = \nabla\varphi(v(x))$

*This should remind us of a projection

We want to minimize this over measure-preserving rearrangements s.

We are looking for measure-preserving s as close as possible to v

→ s is the projection of v onto the set of measure-preserving mappings

- Idea: I can't write the vector field you gave me as the gradient of a convex function, but if I project it onto the right, then I can write it as the gradient of a convex func.

This is what Brenier was focused on.