Measure valued Regression

$$M : \mathbb{P}_2(\mathbb{R}^d)^K \longrightarrow \mathbb{P}_2(\mathbb{R}^d)$$

$$(\mu_1, \dots, \mu_K) \longmapsto \nu$$

M is not an OT map but we will build it using OT.

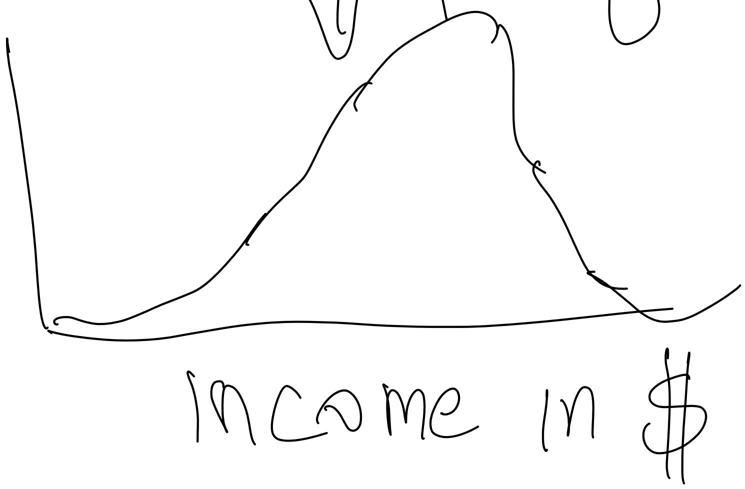
1. Motivation
2. Linearized OT
3. Wasserstein barycenters (signed)
4. Put it together to do measure valued regression

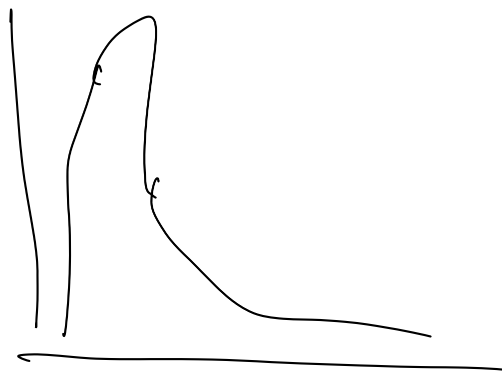
Distributional

Synthetic controls (Gunsilius 2023)

Parallel universe

Policy





We can observe the distribution in our city at times

$t_1, t_2, \dots, t_j > T$

Policy implementation

$\underline{\mu_{t_1}, \mu_{t_2}, \ldots, \mu_{t_j}}, \uparrow, \mu_T$

Policy

We can also observe distributions in other cities $i \in \{1, \ldots, K\}$

$\underline{\mu^i_{t_1}, \mu^i_{t_2}, \ldots, \mu^i_{t_j}}, \underline{\mu^i_T}$

Want to use the μ_t and μ^i_t to learn a map

$$M : \mathbb{P}_2(\mathbb{R}^d)^K \to \mathbb{P}_2(\mathbb{R}^d)$$

for times t before policy implementation

$$\mu^{cf}_T = M(\mu^1_T, \dots, \mu^K_T)$$

Now we can compare μ_T to μ^{cf}_T to determine if the policy had the desired effect.

We want our method to be explainable $\Longrightarrow$ Let's choose a map that is as simple as possible (i.e. a "linear")

First approach to constructing linear maps

- Use linearized OT

Choose a reference measure $\rho \in \mathbb{P}_2(\mathbb{R}^d)$ s.t. ρ is a.c., then by Brenier's thm, given any other measure $\mu \in \mathbb{P}_2(\mathbb{R}^d)$ there exists a <u>unique</u> convex function

$\varphi : \mathbb{R}^d \to \mathbb{R}$ s.t.

$$\nabla\varphi_{\#}\rho = \mu$$

This gives us a correspondance between

$$\mathbb{P}_2(\mathbb{R}^d) \longleftrightarrow \left\{ \varphi : \mathbb{R}^d \to \mathbb{R} : \varphi \text{ is convex and } \int_{\mathbb{R}^d} |\nabla\varphi(x)|^2 d\rho(x) < \infty \right\}$$

The space of convex functions is <u>not</u> a vector space but

It forms a cone in $\mathring{H}^1(\varrho)$

$$= \left\{ f : \mathbb{R}^d \to \mathbb{R} : \int |\nabla f(x)|^2 \, d\varrho(x) < \infty \right\}$$

$$\psi = a_1 \varphi_1 + a_2 \varphi_2$$

Then ψ is a convex function if $a_1, a_2 \geq 0$

Apply linearized OT to

the construction of the map M & our previous setup.

$$\mu^i_{t_1}, \ldots, \mu^i_{t_j}, \mu^i_T$$

$$\nu_{t_1}, \ldots, \nu_{t_j}, \nu_T$$

Given a reference measure ρ we can find OT maps $X^i_{t_1}, \ldots, X^i_{t_j}, X^i_T$

and $Y_{t_1}, \ldots, Y_{t_j}, Y_T$

s.t. $$X_{t_\ell}^{i} {}_\# e = \mu_{t_\ell}^{i}$$

$$Y_{t_\ell} {}_\# e = \nu_{t_\ell}.$$

$$\boxed{\begin{array}{c} X_{t_1}^i, \ldots, X_{t_j}^i, \\ \Big\downarrow \text{linear map} \\ Y_{t_1}, \ldots, Y_{t_j}, \end{array}} \begin{array}{c} X_T^i \\ \\ Y_T \end{array}$$

Now we can use least squares regression to learn map

$$A : (\mathbb{R}^d)^K \to \mathbb{R}^d$$

$$A\begin{pmatrix} X^1 \\ \vdots \\ X^K \end{pmatrix} = Y$$

$$\inf_{A:(\mathbb{R}^d)^K \to \mathbb{R}^d} \sum_{\ell=1}^{j} \left\| \sum_{i=1}^{K} A^i X^i_{t_\ell} - Y_{t_\ell} \right\|^2_{L^2(\rho)}$$

$$\int_{\mathbb{R}^d} \left| A X^i_{t_j}(z) - Y_{t_j}(z) \right|^2 d\rho(z)$$

$$Y = \sum_{i=1}^{K} A^i X^i_T$$

$$Y_\# \varrho = \mu^{cf}_T$$

Without too much computation we get a linear map that sends $\mathbb{P}_2(\mathbb{R}^d)^K \to \mathbb{P}_2(\mathbb{R}^d)$