Distrubitional synthetic controls:

$$\begin{tikzcd}
\mu_{t_1}, \dots, \mu_{t_j} & {} & \mu_T \\
\mu^i_{t_1}, \dots, \mu^i_{t_j} \arrow[u, "\text{train map}"'] & \text{policy} \arrow[u] & , \ \mu^i_T \arrow[d, "M"] \\
& & \mu^{cf}_T
\end{tikzcd}$$

i is indexes different cities $\{1, \dots, K\}$

We want a map

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

We want a simple "linear" map M

IDEA 1: Use linearized OT

choose a reference measure $\sigma \in \mathbb{P}_2(\mathbb{R}^d)$ s.t. σ is a.c.

Reformulate the $\mu^i_{t_j}$ and ν_{t_j} in terms of the maps $X^i_{t_j}$, Y_{t_j} where $X^i_{t_j}$ & Y_{t_j} are the OT maps s.t.

$$X^i_{t_j \#} \ell = \mu^i_{t_j} \qquad Y_{t_j \#} \ell = \nu_{t_j}$$

We construct our map M as follows

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

$$(\mu^1, \ldots, \mu^K) \mapsto Y(\mu^1, \ldots, \mu^K)_{\#} \ell$$

We will construct Y by solving a least squares regression problem

©

$$\min_{A_i} \sum_{\ell=1}^{N} \Big\| \sum_{i=1}^{K} A_i X^i_{t_\ell} - Y_{t_\ell} \Big\|^2_{L^2(e)}$$

$$Y(X^1, \dots, X^K) = \sum_{i=1}^{K} A_i^* X^i$$

$$X^i = \nabla \varphi_i$$ where φ_i is convex

if $Y = \sum_{i=1}^{K} A_i^* \nabla \varphi_i$

when is Y the gradient of a funtion?

For general φ_i this only

happens if $A_i^* = a_i I$

(check by computing $\text{curl}(A \nabla f)$)

If we also want convexity then we need $a_i \geq 0$.

Pros & Cons of Measure valued regression via linearized OT:

Pros

1. Very interpretible
2. "Easy" to compute (just need to solve

several OT problems once +)
some linear algebra

Cons

① Need to choose reference measure ρ.

② Either need to sacrifice expressivity or accept that our map is not an OT map (less guarantees on good behavior).

"Linear" operations on metric spaces + Barycenters

Suppose we have some metric

space (Z, d)

Given some points $z_1, \dots, z_K \in Z$ and coefficients $\lambda_1, \dots, \lambda_K \in [0,1]$ s.t. $\sum_{i=1}^{K} \lambda_i = 1$, then the Barycenter problem seeks to find a point $\bar{z} = \bar{z}(\lambda_1, \dots, \lambda_K)$

$$\bar{z} \in \underset{z \in Z}{\operatorname{argmin}} \sum_{i=1}^{K} \frac{\lambda_i}{2} d(z, z_i)^2$$

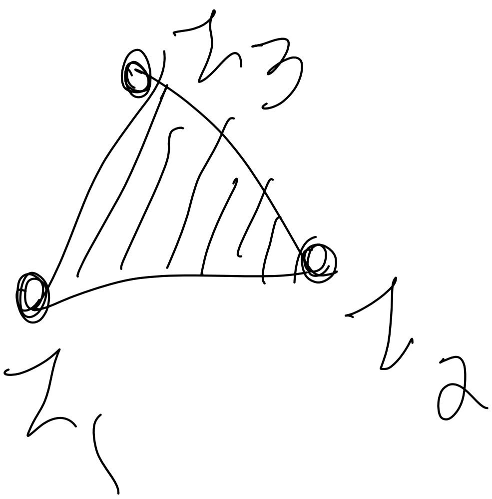
What happens when we do this over $\mathbb{R}^d$

$$\bar{z} \in \underset{z \in \mathbb{R}^d}{\operatorname{argmin}} \sum_{i=1}^{K} \frac{\lambda_i}{2} |z - z_i|^2$$

$$z = \sum_{i=1}^{k} \lambda_i z_i$$

$$\left\{ z(\lambda_1, \ldots, \lambda_k) : \begin{array}{l} \lambda_i \in [0,1] \\ \sum_{i=1}^{k} \lambda_i = 1 \end{array} \right\}$$

= Simplex on $z_1, \ldots, z_k$



$$\left\{ z(\lambda_1, \ldots, \lambda_k) : \begin{array}{l} \sum_{i=1}^{k} \lambda_i = 1 \\ \lambda_i \in \mathbb{R} \end{array} \right\}$$

$= z_1 + \mathrm{span}\{z_2 - z_1, \ldots, z_k - z_1\}$

affine hyperplane containing the simplex on $z_1, \ldots, z_k$

In Euclidean space computing barycenters of points

$\Longleftrightarrow$ linear combinations of points

Over a metric space w/o linear

structure We can view computing barycenters as a synthetic analogue of taking linear combinations.

IDEA: Use Wasserstein barycenter problems to construct "linear" maps between probability spaces.

There is a new challenge of studying these problems w/ negative coefficients.