

img/logo Universidad de Chile

FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA MATEMÁTICA

TÍTULO DE LA MEMORIA/TESIS

TESIS PARA OPTAR AL GRADO DE
MAGÍSTER EN CIENCIAS DE LA INGENIERÍA, MENCIÓN MATEMÁTICAS
APLICADAS

MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL MATEMÁTICO

FRANCISCO MUÑOZ GUAJARDO

PROFESOR GUÍA:
JOAQUÍN FONTBONA TORRES

MIEMBROS DE LA COMISIÓN:
NOMBRE COMPLETO UNO
NOMBRE COMPLETO DOS
NOMBRE COMPLETO TRES

SANTIAGO DE CHILE

2024

Resumen

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Una dedicatoria corta.

Agradecimientos

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.
















Tabla de Contenido

1. Introducción	1
2. Transporte Óptimo de Masas	4
2.1. Notación	4
2.2. El problema de transporte	5
2.3. La Distancia y el Espacio de Wasserstein	7
2.4. El Baricentro de Wasserstein Bayesiano	9
2.4.1. Geodésicas en el Espacio de Wasserstein	10
2.4.2. El Baricentro de Wasserstein Bayesiano	11
Bibliografía	12
Apéndice A. Anexo	13

Índice de Tablas

Índice de Ilustraciones

Lista de tareas pendientes

Un comentario!	2
Otro comentario! Un cambio, otro cambio!	2
 Un comentario	4
 Agregar alguna intro para lo que es una medida de probabilidad? para aquellas personas que vienen de otras áreas?	4
 Se podría dejar notación probabilística?	4
 Se podría agregar una definición de función Lipschitz?	4
 Se podría agregar un ejemplo, el de los cuatro puntos en el plano.	6
 Se podría agregar el típico dibujo en el que se muestra el copling	6
 Tenía pensado en poner el teorema de Brenier (teo 2.1 de [2]), pág 27 para explicar la equivalencia Kantorovich-Monge.	6
 Incluir la formulación de Kantorovich discreto?	6
 Mencionar que esta formulación es más general, y que permite resolver problemas en los que no se puede definir una función de transporte?	7
 Incluir en este punto el teorema de dualidad de Kantorovich-Rubinstein?	7
 Incluir subsections para la distancia, espacio, y convergencia débil?	7
 Se podría poner algún ejemplo con una imagen.	10
 Agregar la referencia del Lecture	10
 Agregar en la def que estos se llamarán planes de transporte	11
 Comentar que aquí se está tomando ideas de la tesis de Gonzalo Ríos	11

Capítulo 2

Transporte Óptimo de Masas

En este capítulo se abordará el problema de transporte óptimo, la distancia de Wasserstein, y el problema de los baricentros de Wasserstein. Además, se presentarán algunas propiedades de la distancia de Wasserstein, las cuales serán de utilidad en el desarrollo de este trabajo. La notación y definiciones utilizadas en este capítulo se encuentran basadas en [3] y [2]. Sin embargo, antes de empezar a enunciar definiciones y propiedades, se sentarán la notación y definiciones básicas que se utilizarán a lo largo de este trabajo.

Joaquín

Un comentario

2.1. Notación

Francisco: Agregar alguna intro para lo que es una medida de probabilidad? para aquellas personas que vienen de otras áreas?

Definición 2.1.1.

Francisco: Se podría dejar notación probabilística?

Se definen los siguientes espacios:

- $(\mathcal{X}, \text{dist})$ es un espacio Polaco, si \mathcal{X} es un espacio métrico, completo y separable.
- $\mathcal{P}(\mathcal{X})$ denotará al conjunto de medidas de probabilidad en \mathcal{X} , utilizando la σ -álgebra de Borel.
- $\mathcal{P}_{\text{ac}}(\mathcal{X})$ denotará al conjunto de medidas de probabilidad absolutamente continuas con respecto a una medida de referencia λ (como por ejemplo, la de Lebesgue o la cuenta puntos), utilizando la σ -álgebra de Borel.
- $\mathcal{C}(\mathcal{X})$ denotará al conjunto de funciones continuas en \mathcal{X} .
- $\text{Lip}_k(\mathcal{X})$ denotará al conjunto de funciones k -Lipschitz en \mathcal{X} . Mientras que se asumirá que $\text{Lip}(\mathcal{X})$ denotará al conjunto de funciones 1-Lipschitz en \mathcal{X} .

Francisco: Se podría agregar una definición de función Lipschitz?

Definición 2.1.2. Se definirá el *simplex* de dimension n como el conjunto de vectores de \mathbb{R}^n cuyas componentes suman 1, es decir,

$$\Sigma_n \stackrel{\text{def}}{=} \left\{ x \in [0, 1]^n : \sum_{i=1}^n x_i = 1 \right\}, \quad (2.1)$$

y a los elementos pertenecientes al simplex se les llamará *vectores de probabilidad*.

Definición 2.1.3. Dados $\mu \in \mathcal{P}(\mathcal{X})$ y $\nu \in \mathcal{P}(\mathcal{Y})$, se denotará por $\text{Cpl}(\mu, \nu)$ al conjunto de medidas de probabilidad en $\mathcal{X} \times \mathcal{Y}$ cuyas proyecciones marginales sean μ y ν , es decir,

$$\text{Cpl}(\mu, \nu) \stackrel{\text{def}}{=} \{ \gamma \in \mathcal{P}(\mathcal{X} \times \mathcal{Y}) : \gamma(A \times \mathcal{Y}) = \mu(A), \gamma(\mathcal{X} \times B) = \nu(B), \forall A \subseteq \mathcal{X}, B \subseteq \mathcal{Y} \}. \quad (2.2)$$

Definición 2.1.4. Para una función medible $T : \mathcal{X} \rightarrow \mathcal{Y}$ se define el *operador push-forward* de T como la aplicación $T_{\#} : \mathcal{P}(\mathcal{X}) \rightarrow \mathcal{P}(\mathcal{Y})$ que satisface la siguiente relación:

$$\int_{\mathcal{X}} f(x) dT_{\#}\mu(x) = \int_{\mathcal{Y}} f(T(x)) d\mu(x), \quad \forall f \in \mathcal{C}(\mathcal{Y}), \quad (2.3)$$

para toda $\mu \in \mathcal{P}(\mathcal{X})$. Adicionalmente, el operador push-forward se puede definir como aquel operador que satisface la siguiente relación:

$$\forall A \subseteq \mathcal{Y} \text{ medible}, \quad T_{\#}\mu(A) = \mu(T^{-1}(A)). \quad (2.4)$$

Observación 2.1.5. Se puede notar que $T_{\#}$ preserva la positividad y la masa total, es decir, si $\mu \in \mathcal{P}(\mathcal{X})$, entonces $T_{\#}\mu \in \mathcal{P}(\mathcal{Y})$.

Observación 2.1.6. Para el caso en que la medida $\mu \in \mathcal{P}(\mathcal{X})$ sea una medida discreta¹, entonces el operador $T_{\#}$ lo que hará será intercambiar la masa de cada punto de \mathcal{X} a su imagen en \mathcal{Y} , es decir,

$$T_{\#}\mu = \sum_{i=0}^n m_i \delta_{T(x_i)}. \quad (2.5)$$

2.2. El problema de transporte

El problema de Monge

En esta sección se introducirá el problema de transporte óptimo de Monge. En este problema se busca transportar la masa de la medida $\mu \in \mathcal{P}(\mathcal{X})$ a la medida $\nu \in \mathcal{P}(\mathcal{Y})$ a través de una función medible $T : \mathcal{X} \rightarrow \mathcal{Y}$, a la que llamaremos como *función de transporte* o *mapa de transporte*, utilizando una *función de coste* $c(x, y)$, que representa el costo de transportar la masa del punto x al punto y . Más formalmente, este problema se puede definir de la siguiente manera:

¹i.e. $\mu = \sum_{i=0}^n m_i \delta_{x_i}$ con $m \in \Sigma_n$, $x_1, \dots, x_n \in \mathcal{X}$ y δ_x la medida de Dirac en x

Definición 2.2.1. Dados $\mu \in \mathcal{P}(\mathcal{X})$ y $\nu \in \mathcal{P}(\mathcal{Y})$, se define el *problema de transporte óptimo* como el problema de encontrar una función de transporte $T : \mathcal{X} \rightarrow \mathcal{Y}$ que minimice el costo total de transporte, es decir, que minimice la siguiente expresión:

$$\inf_{T: T_{\#}\mu=\nu} \int_{\mathcal{X}} c(x, T(x)) d\mu(x), \quad (2.6)$$

Observación 2.2.2. Cuando las medidas μ y ν son discretas, es decir, se representan de la siguiente manera:

$$\mu = \sum_{i=0}^n m_i \delta_{x_i}, \quad \nu = \sum_{j=0}^m n_j \delta_{y_j}, \quad (2.7)$$

donde $m \in \Sigma_n$, $n \in \Sigma_m$, $x_i \in \mathcal{X}$, $y_j \in \mathcal{Y}$, entonces el problema de transporte óptimo se puede representar de la siguiente manera:

$$\inf_{T: T(x_i)=y_j} \sum_{i=0}^n c(x_i, T(x_i)) m_i. \quad (2.8)$$

Cabe destacar que este problema no siempre tiene solución (generalmente no la tiene si $m > n$), y que en caso de tenerla, no siempre es única.

Francisco: Se podría agregar un ejemplo, el de los cuatro puntos en el plano.

El problema de Kantorovich

Como se mencionó en la Observación 2.2.2, el problema de transporte óptimo de Monge no siempre tiene solución, y en caso de tenerla, no siempre es única. Motivado por esto, Kantorovich propuso una formulación alternativa del problema de transporte óptimo, que si tiene solución (aunque puede que no sea única). Este problema se puede definir de la siguiente manera:

Definición 2.2.3. Dados $\mu \in \mathcal{P}(\mathcal{X})$ y $\nu \in \mathcal{P}(\mathcal{Y})$, se define el *problema de transporte óptimo de Kantorovich* como el problema de encontrar una medida de probabilidad $\gamma \in \text{Cpl}(\mu, \nu)$ que minimice el costo total de transporte, es decir, que minimice la siguiente expresión:

$$\inf_{\gamma \in \text{Cpl}(\mu, \nu)} \int_{\mathcal{X} \times \mathcal{Y}} c(x, y) d\gamma(x, y). \quad (2.9)$$

Al conjunto de medidas de probabilidad que resuelven este problema se le llama *transportes óptimos*.

Francisco: Se podría agregar el típico dibujo en el que se muestra el coupling

Francisco: Tenía pensado en poner el teorema de Brenier (teo 2.1 de [2]), pág 27 para explicar la equivalencia Kantorovich-Monge.

Francisco: Incluir la formulación de Kantorovich discreto?

Francisco: Mencionar que esta formulación es más general, y que permite resolver problemas en los que no se puede definir una función de transporte?

Francisco: Incluir en este punto el teorema de dualidad de Kantorovich-Rubinstein?

2.3. La Distancia y el Espacio de Wasserstein

Francisco

Incluir subsections para la distancia, espacio, y convergencia débil?

En esta sección se demostrará que, al evaluar la expresión (2.9) para una función de coste con distancia, se obtiene una distancia entre medidas de probabilidad. Revisaremos algunas propiedades de esta distancia, para concluir que esta distancia metriza la convergencia débil entre medidas de probabilidad.

Definición 2.3.1 (La distancia de Wasserstein). Sea $(\mathcal{X}, \text{dist})$ un espacio Polaco y sea $p \geq 1$. Para dos medidas μ, ν sobre \mathcal{X} , la distancia de Wasserstein de orden p entre μ y ν es definida por medio de la fórmula

$$W_p(\mu, \nu) \stackrel{\text{def}}{=} \left(\inf_{\gamma \in \text{Cpl}(\mu, \nu)} \int_{\mathcal{X} \times \mathcal{X}} \text{dist}(x, y)^p d\gamma(x, y) \right)^{\frac{1}{p}}. \quad (2.10)$$

Ejemplo 2.3.2. $W_1(\delta_x, \delta_y) = \text{dist}(x, y)$. Notemos que en este ejemplo, se puede interpretar que la distancia de Wasserstein metriza el “esfuerzo” de llevar la masa del punto x al punto y .

Notemos que, en estricto rigor, W_p no es una distancia en sí, dado que puede tomar valores de $+\infty$, sin embargo, se puede demostrar que W_p satisface los axiomas de ser una distancia. No se demostrará este hecho, pero se puede encontrar una demostración en [3], pág 94.

Por tanto, resulta natural definir el espacio en el que la distancia de Wasserstein tome valores finitos.

Definición 2.3.3 (El espacio de Wasserstein). Con los mismos supuestos que en la Definición 2.3.1, se define el espacio de Wasserstein de orden p por medio de

$$\mathcal{W}_p(\mathcal{X}) \stackrel{\text{def}}{=} \left\{ \mu \in \mathcal{P}(\mathcal{X}) : \int_{\mathcal{X}} \text{dist}(x, x_0)^p d\mu(x) < \infty \right\}, \quad (2.11)$$

donde $x_0 \in \mathcal{X}$ es un punto fijo arbitrario. De esta forma, W_p define una distancia (finita) sobre $\mathcal{W}_p(\mathcal{X})$.

En palabras simples, el espacio de Wasserstein de orden p es el conjunto de medidas de probabilidad en \mathcal{X} cuyo momento de orden p es finito. Lo interesante del espacio de Wasserstein, es que su respectiva distancia lo metriza, como lo dice el siguiente teorema:

Teorema 2.3.4. Si $(\mathcal{X}, \text{dist})$ es un espacio Polaco, entonces el espacio de Wasserstein $\mathcal{W}_p(\mathcal{X})$, metrizado por la distancia de Wasserstein W_p , es también un espacio Polaco.

DEMOSTRACIÓN. Revisar la demostración del Teorema 6.18 en [3, p. 105]

□

A partir de ahora, se asumirá que el espacio $\mathcal{W}_p(\mathcal{X})$ siempre estará equipado con su respectiva distancia W_p .

Observación 2.3.5. A través de la desigualdad de Hölder, se puede demostrar que para $p \leq q$, se tiene que $W_p(\mu, \nu) \leq W_q(\mu, \nu)$, para toda $\mu, \nu \in \mathcal{W}_p(\mathcal{X})$. Y por tanto, las topologías inducidas por las distancias de Wasserstein se van encajonando.

En particular, la distancia de Wasserstein de orden 1, es la más débil de todas. Como norma general, la distancia W_1 es la más flexible y fácil de acotar, mientras que la distancia W_2 posee mejores propiedades geométricas, pero es más difícil de trabajar.

Vista la distancia y el espacio de Wasserstein, se presentará una caracterización de convergencia en este espacio. Para ello, se definirá la convergencia débil entre medidas de probabilidad.

Definición 2.3.6 (Convergencia Débil). Sea $(\mathcal{X}, \text{dist})$ un espacio Polaco y sea $p \geq 1$. Se dice que una sucesión de medidas de probabilidad $(\mu_n)_{n \in \mathbb{N}} \subset \mathcal{W}_p(\mathcal{X})$ converge débilmente a $\mu \in \mathcal{W}_p(\mathcal{X})$ si

$$\forall \varphi \in \mathcal{C}_b(\mathcal{X}), \quad \int_{\mathcal{X}} \varphi(x) d\mu_n(x) \rightarrow \int_{\mathcal{X}} \varphi(x) d\mu(x). \quad (2.12)$$

y lo denotaremos por $\mu_n \Rightarrow \mu$.

Nota 2.3.7. Intuitivamente, que una sucesión de medidas de probabilidad converjan débilmente a una medida μ significa que es la forma “más fácil” que tiene la sucesión de converger a μ .

Teorema 2.3.8 (La Distancia de Wasserstein Metriza la Convergencia Débil). Sea $(\mathcal{X}, \text{dist})$ un espacio Polaco y sea $p \geq 1$. Entonces, la distancia de Wasserstein W_p metriza la convergencia débil en $\mathcal{W}_p(\mathcal{X})$.

Observación 2.3.9. En otras palabras, si $(\mu_n)_{n \in \mathbb{N}}$ es una sucesión de medidas de probabilidad en $\mathcal{W}_p(\mathcal{X})$ y $\mu \in \mathcal{W}_p(\mathcal{X})$ otra medida, entonces $\mu_n \Rightarrow \mu$ si y sólo si $W_p(\mu_n, \mu) \rightarrow 0$.

Ejemplo 2.3.10. Consideremos las siguientes distancias y divergencias entre medidas de probabilidad:

$$\begin{aligned} \text{TV}(\mu, \nu) &\stackrel{\text{def}}{=} \sup_{A \subseteq \mathcal{X}} |\mu(A) - \nu(A)|, \\ \text{KL}(\mu \mid \nu) &\stackrel{\text{def}}{=} \int_{\mathcal{X}} \log \left(\frac{d\mu}{d\nu} \right) d\mu(x), \\ \text{JS}(\mu, \nu) &\stackrel{\text{def}}{=} \text{KL} \left(\mu \mid \frac{\mu + \nu}{2} \right) + \text{KL} \left(\nu \mid \frac{\mu + \nu}{2} \right), \end{aligned}$$

donde la primera es la distancia total variación, la segunda es la divergencia de Kullback-Leibler, y la tercera es la divergencia de Jensen-Shannon.

Si consideramos δ_θ y δ_0 medidas de Dirac centradas en θ y 0 respectivamente, entonces se puede demostrar que

$$\begin{aligned} W_1(\delta_\theta, \delta_0) &= |\theta| & \text{TV}(\delta_\theta, \delta_0) &= \begin{cases} 1 & \text{si } \theta \neq 0 \\ 0 & \text{si } \theta = 0 \end{cases} \\ \text{KL}(\delta_\theta \mid \delta_0) &= \begin{cases} +\infty & \text{si } \theta \neq 0 \\ 0 & \text{si } \theta = 0 \end{cases} & \text{JS}(\delta_\theta, \delta_0) &= \begin{cases} \log(2) & \text{si } \theta \neq 0 \\ 0 & \text{si } \theta = 0 \end{cases} \end{aligned}$$

Entonces, si tomamos $\theta = \frac{1}{n}$ y dejamos que $n \rightarrow \infty$, se tiene que $W_1(\delta_\theta, \delta_0) \rightarrow 0$, pero el resto de distancias y divergencias no convergen a 0. Por tanto, se puede notar que la distancia de Wasserstein es la única que es capaz de distinguir entre medidas de probabilidad que no tienen soporte en el mismo punto, gracias a que metriza la convergencia débil.

2.4. El Baricentro de Wasserstein Bayesiano

La Media de Fréchet

En esta sección se revisará el concepto de media de Fréchet, el cual es una generalización de la noción de promedio para espacios métricos. Este concepto será clave para definir el baricentro de Wasserstein.

Definición 2.4.1 (Funcional y Media de Fréchet). Sea $(\mathcal{X}, \text{dist})$ un espacio Polaco. Sean x_1, \dots, x_n puntos en \mathcal{X} y sean $w_1, \dots, w_n \in \mathbb{R}$ pesos asociados a los puntos. Para cada $p \in \mathcal{X}$, se define el *funcional de Fréchet* por

$$\Psi(p) \stackrel{\text{def}}{=} \sum_{i=1}^n w_i \text{dist}(p, x_i)^2. \quad (2.13)$$

Y, en caso de que exista un punto $m \in \mathcal{X}$ que minimice el funcional Ψ , entonces este se definirá como la *media de Fréchet* de los puntos x_1, \dots, x_n . Es decir, es aquel punto tal que minimiza el siguiente problema:

$$m \stackrel{\text{def}}{=} \arg \min_{p \in \mathcal{X}} \sum_{i=1}^n w_i \text{dist}(p, x_i)^2. \quad (2.14)$$

Ejemplo 2.4.2. Tomemos $x_1, x_2, x_3 \in \mathbb{R}^2$ tres puntos en el plano, formando un triángulo. Si se define el promedio (o el *baricentro*, en el contexto de un triángulo) de estos puntos por $\bar{x} = \frac{1}{3}(x_1 + x_2 + x_3)$, entonces se puede comprobar fácilmente que este es el único que minimiza el funcional de Fréchet:

$$F(p) = \frac{1}{3} \sum_{i=1}^3 \|p - x_i\|^2. \quad (2.15)$$

Dado que este funcional se puede descomponer de la siguiente manera:

$$F(p) = F(\bar{x}) + \|p - \bar{x}\|^2, \quad (2.16)$$

se puede ver que la media de Fréchet generaliza la noción de promedio.

Observación 2.4.3. La razón por la que resulta interesante estudiar este concepto, es que sólo utiliza nociones métricas, y se desliga de la noción vectorial. Como se vió en el Ejemplo 2.4.2, el promedio utilizó nociones vectoriales (suma, ponderación) mientras que la media de Fréchet utilizó nociones métricas, resultando en el mismo promedio.

Sin embargo, el hecho de que se pueda cambiar la distancia, hace que el baricentro cambie, y dependa de ésta.

El Baricentro de Wasserstein

Como se vió en la sección anterior, la media de Fréchet permite definir una noción de promedio, en espacios métricos. El Teorema 2.3.4 nos dice que $(\mathcal{W}_p(\mathcal{X}), W_p)$ es un espacio métrico, y por tanto, se puede definir su respectivo “promedio”:

Definición 2.4.4. Sean $\mu_1, \dots, \mu_n \in \mathcal{W}_p(\mathcal{X})$ y sean $w = (w_1, \dots, w_n) \in \Sigma_n$ sus pesos asociados. El *baricentro de Wasserstein* se define por medio de

$$\bar{\mu} \stackrel{\text{def}}{=} \arg \inf_{\nu \in \mathcal{W}_p(\mathcal{X})} \sum_{i=1}^n w_i W_p(\nu, \mu_i)^p \quad (2.17)$$

Francisco: Se podría poner algún ejemplo con una imagen.

Es posible generalizar aún más la noción de baricentro de Wasserstein a una colección infinita de medidas. Esto se puede hacer considerando una medida $\Gamma \in \mathcal{P}(\mathcal{P}(\mathcal{X}))$, que cumplirá el rol de los pesos w_1, \dots, w_n en la definición anterior. Esto se puede formalizar en la siguiente definición:

Definición 2.4.5. Sea $\Gamma \in \mathcal{P}(\mathcal{P}(\mathcal{X}))$ una medida. El baricentro de Wasserstein se puede (re)-definir como aquel que minimice el siguiente problema:

$$\bar{\mu} \stackrel{\text{def}}{=} \arg \inf_{\mu \in \mathcal{P}(\mathcal{X})} \int_{\mathcal{P}(\mathcal{X})} W_p(\mu, \nu)^p d\Gamma(\nu) \quad (2.18)$$

2.4.1. Geodésicas en el Espacio de Wasserstein

En esta sección, se utilizará la distancia de Wasserstein para definir una noción de geodésica en el espacio de Wasserstein. Esta noción será de utilidad para definir el baricentro de Wasserstein Bayesiano.

Francisco: Agregar la referencia del Lecture

Definición 2.4.6. Sea $(\mathcal{X}, \text{dist})$ un espacio métrico. Una *geodésica a velocidad constante* entre dos puntos, $x_0, x_1 \in \mathcal{X}$ es una curva continua $x : [0, 1] \rightarrow \mathcal{X}$ tal que para cada $s, t \in [0, 1]$, $\text{dist}(x_s, x_t) = |s - t| \text{dist}(x_0, x_1)$.

Proposición 2.4.7. Sean $\mu_0, \mu_1 \in \mathcal{W}_p(\mathcal{X})$, con $\mathcal{X} \in \mathbb{R}^d$ compacto y convexo. Sea $\gamma \in \text{Cpl}(\mu_0, \mu_1)$ un plan de transporte óptimo. Definamos

$$\mu_t \stackrel{\text{def}}{=} (\pi_t)_\# \gamma, \quad \text{donde } \pi_t(x, y) = (1 - t)x + ty. \quad (2.19)$$

Entonces, la curva μ_t es una geodésica a velocidad constante entre μ_0 y μ_1 .

Ejemplo 2.4.8. Si existe un mapa de transporte óptimo T entre μ_0 y μ_1 , entonces la geodésica definida arriba es $\mu_t = ((1 - t)\text{id} + tT)_\# \mu_0$.

Observación 2.4.9. En caso en que $\mu_0 = \sum_{i=1}^n m_i \delta_{x_i}$ y $\mu_1 = \sum_{i=1}^n m_i \delta_{y_i}$, con $m \in \Sigma_n$ y $x_1, \dots, x_n, y_1, \dots, y_n \in \mathcal{X}$, entonces la geodésica definida arriba se puede escribir de la siguiente manera:

$$\mu_t = \sum_{i=1}^n m_i \delta_{(1-t)x_i + ty_i}, \quad (2.20)$$

lo que se puede interpretar como que la masa m_i del punto x_i se transporta al punto y_i a velocidad constante.

2.4.2. El Baricentro de Wasserstein Bayesiano

Francisco: Comentar que aquí se está tomando ideas de la tesis de Gonzalo Ríos

Consideremos muestras $\mathcal{D} = \{x_1, \dots, x_n\}$ en un espacio \mathcal{X} y un conjunto de modelos factibles o medidas de probabilidad $\mathcal{M} \subseteq \mathcal{P}(\mathcal{X})$. Aprender un modelo, también conocido como la *selección de un modelo*, de \mathcal{D} consiste en escoger un elemento $\mu \in \mathcal{M}$ que mejor explique los datos, si es que estos hubieran sido generados por μ , dado algún criterio.

Se considerará una medida de probabilidad $\Pi \in \mathcal{P}(\mathcal{M})$ sobre un espacio de modelos \mathcal{M} un *prior* fijo.

Francisco

Agregar en la def que estos se llamarán planes de transporte

Bibliografía

- [1] Julio Backhoff-Veraguas, Joaquin Fontbona, Gonzalo Rios, and Felipe Tobar. Bayesian learning with wasserstein barycenters, 2018.
- [2] Gabriel Peyré and Marco Cuturi. Computational optimal transport. *Foundations and Trends in Machine Learning*, 11(5-6):355–607, 2019.
- [3] Cédric Villani. *Optimal Transport: Old and New*, volume 338. Springer Berlin Heidelberg, 2009.

Apéndice A

Anexo

Quisque facilisis auctor sapien. Pellentesque gravida hendrerit lectus. Mauris rutrum sodales sapien. Fusce hendrerit sem vel lorem. Integer pellentesque massa vel augue. Integer elit tortor, feugiat quis, sagittis et, ornare non, lacus. Vestibulum posuere pellentesque eros. Quisque venenatis ipsum dictum nulla. Aliquam quis quam non metus eleifend interdum. Nam eget sapien ac mauris malesuada adipiscing. Etiam eleifend neque sed quam. Nulla facilisi. Proin a ligula. Sed id dui eu nibh egestas tincidunt. Suspendisse arcu.

Maecenas dui. Aliquam volutpat auctor lorem. Cras placerat est vitae lectus. Curabitur massa lectus, rutrum euismod, dignissim ut, dapibus a, odio. Ut eros erat, vulputate ut, interdum non, porta eu, erat. Cras fermentum, felis in porta congue, velit leo facilisis odio, vitae consectetur lorem quam vitae orci. Sed ultrices, pede eu placerat auctor, ante ligula rutrum tellus, vel posuere nibh lacus nec nibh. Maecenas laoreet dolor at enim. Donec molestie dolor nec metus. Vestibulum libero. Sed quis erat. Sed tristique. Duis pede leo, fermentum quis, consectetur eget, vulputate sit amet, erat.

Donec vitae velit. Suspendisse porta fermentum mauris. Ut vel nunc non mauris pharetra varius. Duis consequat libero quis urna. Maecenas at ante. Vivamus varius, wisi sed egestas tristique, odio wisi luctus nulla, lobortis dictum dolor ligula in lacus. Vivamus aliquam, urna sed interdum porttitor, metus orci interdum odio, sit amet euismod lectus felis et leo. Praesent ac wisi. Nam suscipit vestibulum sem. Praesent eu ipsum vitae pede cursus venenatis. Duis sed odio. Vestibulum eleifend. Nulla ut massa. Proin rutrum mattis sapien. Curabitur dictum gravida ante.

Phasellus placerat vulputate quam. Maecenas at tellus. Pellentesque neque diam, dignissim ac, venenatis vitae, consequat ut, lacus. Nam nibh. Vestibulum fringilla arcu mollis arcu. Sed et turpis. Donec sem tellus, volutpat et, varius eu, commodo sed, lectus. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Quisque enim arcu, suscipit nec, tempus at, imperdiet vel, metus. Morbi volutpat purus at erat. Donec dignissim, sem id semper tempus, nibh massa eleifend turpis, sed pellentesque wisi purus sed libero. Nullam lobortis tortor vel risus. Pellentesque consequat nulla eu tellus. Donec velit. Aliquam fermentum, wisi ac rhoncus iaculis, tellus nunc malesuada orci, quis volutpat dui magna id mi. Nunc vel ante. Duis vitae lacus. Cras nec ipsum.

Morbi nunc. Aliquam consectetur varius nulla. Phasellus eros. Cras dapibus porttitor risus. Maecenas ultrices mi sed diam. Praesent gravida velit at elit vehicula porttitor. Phasellus nisl mi, sagittis ac, pulvinar id, gravida sit amet, erat. Vestibulum est. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur id sem elementum leo rutrum hendrerit. Ut at mi. Donec tincidunt faucibus massa. Sed turpis quam, sollicitudin a, hendrerit eget, pretium ut, nisl. Duis hendrerit ligula. Nunc pulvinar congue urna.

Nunc velit. Nullam elit sapien, eleifend eu, commodo nec, semper sit amet, elit. Nulla lectus risus, condimentum ut, laoreet eget, viverra nec, odio. Proin lobortis. Curabitur dictum arcu vel wisi. Cras id nulla venenatis tortor congue ultrices. Pellentesque eget pede. Sed eleifend sagittis elit. Nam sed tellus sit amet lectus ullamcorper tristique. Mauris enim sem, tristique eu, accumsan at, scelerisque vulputate, neque. Quisque lacus. Donec et ipsum sit amet elit nonummy aliquet. Sed viverra nisl at sem. Nam diam. Mauris ut dolor. Curabitur ornare tortor cursus velit.

Morbi tincidunt posuere arcu. Cras venenatis est vitae dolor. Vivamus scelerisque semper mi. Donec ipsum arcu, consequat scelerisque, viverra id, dictum at, metus. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut pede sem, tempus ut, porttitor bibendum, molestie eu, elit. Suspendisse potenti. Sed id lectus sit amet purus faucibus vehicula. Praesent sed sem non dui pharetra interdum. Nam viverra ultrices magna.

Aenean laoreet aliquam orci. Nunc interdum elementum urna. Quisque erat. Nullam tempor neque. Maecenas velit nibh, scelerisque a, consequat ut, viverra in, enim. Duis magna. Donec odio neque, tristique et, tincidunt eu, rhoncus ac, nunc. Mauris malesuada malesuada elit. Etiam lacus mauris, pretium vel, blandit in, ultricies id, libero. Phasellus bibendum erat ut diam. In congue imperdiet lectus.

Aenean scelerisque. Fusce pretium porttitor lorem. In hac habitasse platea dictumst. Nulla sit amet nisl at sapien egestas pretium. Nunc non tellus. Vivamus aliquet. Nam adipiscing euismod dolor. Aliquam erat volutpat. Nulla ut ipsum. Quisque tincidunt auctor augue. Nunc imperdiet ipsum eget elit. Aliquam quam leo, consectetur non, ornare sit amet, tristique quis, felis. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Pellentesque interdum quam sit amet mi. Pellentesque mauris dui, dictum a, adipiscing ac, fermentum sit amet, lorem.

Ut quis wisi. Praesent quis massa. Vivamus egestas risus eget lacus. Nunc tincidunt, risus quis bibendum facilisis, lorem purus rutrum neque, nec porta tortor urna quis orci. Aenean aliquet, libero semper volutpat luctus, pede erat lacinia augue, quis rutrum sem ipsum sit amet pede. Vestibulum aliquet, nibh sed iaculis sagittis, odio dolor blandit augue, eget mollis urna tellus id tellus. Aenean aliquet aliquam nunc. Nulla ultricies justo eget orci. Phasellus tristique fermentum leo. Sed massa metus, sagittis ut, semper ut, pharetra vel, erat. Aliquam quam turpis, egestas vel, elementum in, egestas sit amet, lorem. Duis convallis, wisi sit amet mollis molestie, libero mauris porta dui, vitae aliquam arcu turpis ac sem. Aliquam aliquet dapibus metus.

Vivamus commodo eros eleifend dui. Vestibulum in leo eu erat tristique mattis. Cras at elit. Cras pellentesque. Nullam id lacus sit amet libero aliquet hendrerit. Proin placerat, mi non elementum laoreet, eros elit tincidunt magna, a rhoncus sem arcu id odio. Nulla eget leo a

leo egestas facilisis. Curabitur quis velit. Phasellus aliquam, tortor nec ornare rhoncus, purus urna posuere velit, et commodo risus tellus quis tellus. Vivamus leo turpis, tempus sit amet, tristique vitae, laoreet quis, odio. Proin scelerisque bibendum ipsum. Etiam nisl. Praesent vel dolor. Pellentesque vel magna. Curabitur urna. Vivamus congue urna in velit. Etiam ullamcorper elementum dui. Praesent non urna. Sed placerat quam non mi. Pellentesque diam magna, ultricies eget, ultrices placerat, adipiscing rutrum, sem.