



UNIVERSIDAD NACIONAL DEL LITORAL
FACULTAD DE INGENIERÍA QUÍMICA
INSTITUTO DE MATEMÁTICA APLICADA DEL LITORAL

Tesis presentada como parte de los requisitos de la Universidad Nacional del Litoral
para la obtención del Grado Académico de

Doctor en Matemática

En el campo del **Análisis Matemático**

Título

**Distancias entre espacios métricos con medidas.
Concentración de grafos aleatorios ponderados y con
atributos. Aplicaciones**

Autor

Carlos Exequiel Arias

Institución donde se realizó la investigación:

Instituto de Matemática Aplicada del Litoral Dra. "Eleonor Harboure"

CONICET – UNL

Director de Tesis: **Hugo Aimar**

Codirector de Tesis: **Diego Tomassi**

Jurado compuesto por

J. Fernández Bonder, P. Groisman, R. Spies

Año de presentación: 2024

*A mi madre Isabel,
por ser el pilar de mi vida
y mi más grande inspiración.*

Agradecimientos

A mis directores, el Dr. Hugo Aimar y el Dr. Diego Tomassi, les agradezco por la confianza que depositaron en mí. A la Dra. Ivana Gómez, por su especial acompañamiento y guía en este trabajo, ya que su aporte fue fundamental. A los tres, por brindarme la oportunidad de formar parte de sus proyectos y por guiarme en esta etapa de aprendizaje, lo cual ha contribuido no solo a mi desarrollo profesional, sino también a mi crecimiento personal.

Al Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), por haber financiado mi estudio. Al Instituto de Matemática Aplicada del Litoral (IMAL) por haberme proporcionado un lugar de trabajo cómodo y confortable. A mis compañeros del IMAL, por su compañerismo y apoyo brindado durante estos años, lo que ha facilitado mi estadía.

A mi familia y amigos, por su amor y apoyo incondicional, a sus cálidas reuniones cada vez que regresaba y por su constante aliento en cada paso que daba, han sido un aliciente invaluable. A Isaías, por habernos conocido lejos de casa, demostrando ser un gran compañero y amigo, compartiendo juntos esta experiencia. Gracias a todos ustedes por brindarme un empujón, cada uno a su manera.

A mis padres, quienes me incentivaron a formarme y por brindarme sostén a lo largo de la vida, en especial a mi madre Isabel, quien me alentó a perseguir y cumplir un sueño, superándome día a día. A mis hermanas, Elizabeth y Cinthia, quienes me han acompañado en el proceso, brindándome su apoyo constante, escuchándome y aconsejándome, fueron un pilar fundamental. También, quiero agradecer a mi compañera de la vida, Griselda, por estar presente en esta etapa y por su gran apoyo. Este logro es tanto mío como de ustedes.

Esta meta se volvió posible gracias a todos ustedes. De corazón, les agradezco profundamente.

Índice de contenidos

| | |
|--|-----|
| | III |
| Agradecimientos | V |
| Resumen | V |
| Introducción de la Tesis | VII |
| Capítulo 1. Grafos no dirigidos y ponderados | 1 |
| 1.1. Introducción | 1 |
| 1.2. Energía de funciones reales definidas en los vértices de grafos no dirigidos ponderados. Ecuación de Euler-Lagrange | 2 |
| 1.3. El operador de Laplace en grafos ponderados no dirigidos con atributos | 5 |
| 1.4. Ejemplos de grafos clásicos y sus espectros | 10 |
| 1.5. Un grafo para el transporte público en AMBA | 15 |
| Capítulo 2. Atracción, afinidad, topología y distancia | 19 |
| 2.1. Introducción | 19 |
| 2.2. Topología y sistemas de entornos | 19 |
| 2.3. Topologías inducidas por protométricas | 21 |
| 2.4. Topologías inducidas por afinidades en conjuntos arbitrarios | 22 |
| 2.5. Métricas, casi-métricas y pseudométricas | 23 |
| 2.6. Metrización de afinidades transitivas | 24 |
| 2.7. Métricas Difusivas | 26 |
| Capítulo 3. Dinámica topológica por umbralado de afinidades | 27 |
| 3.1. Introducción | 27 |
| 3.2. Umbrales de una afinidad y la familia de topologías inducidas | 28 |
| 3.3. Algunos ejemplos sencillos | 29 |
| 3.4. Conectividad topológica y conectividad por caminos de grafos | 35 |

| | | |
|-------------|---|----|
| 3.5. | Dinámica de la conectividad en el transporte público en AMBA | 38 |
| Capítulo 4. | Distancia y Afinidad de Hausdorff | 43 |
| 4.1. | Introducción | 43 |
| 4.2. | Un algoritmo en Python para el cálculo de d_H entre subconjuntos finitos de un espacio métrico | 45 |
| 4.3. | Afinidad de Hausdorff entre subconjuntos de un espacio con afinidad | 49 |
| 4.4. | Un algoritmo en Python para el cálculo de W_H entre subconjuntos finitos de un espacio con afinidad | 50 |
| 4.5. | Casi-métrica de Hausdorff | 53 |
| 4.6. | Afinidades transitivas y casi-métricas de Hausdorff | 54 |
| Capítulo 5. | Una distancia entre grafos métricos. El enfoque de Gromov-Lipschitz | 57 |
| 5.1. | Introducción | 57 |
| 5.2. | La métrica de Gromov-Lipschitz | 58 |
| 5.3. | Metrización de Gromov-Lipschitz de una clase de grafos métricos | 61 |
| 5.4. | Algunos ejemplos | 68 |
| 5.5. | Distancia de Gromov-Lipschitz entre métricas difusivas en un grafo. Aplicación al grafo SUBE en AMBA | 73 |
| Capítulo 6. | Distancia entre grafos métricos. El enfoque Gromov-Hausdorff | 77 |
| 6.1. | Introducción | 77 |
| 6.2. | La métrica de Gromov-Hausdorff | 77 |
| 6.3. | Metrización de Gromov-Hausdorff de una clase de grafos métricos | 81 |
| 6.4. | Algunos ejemplos | 82 |
| 6.5. | Cálculo de distancia de Gromov-Hausdorff usando correspondencias | 85 |
| 6.6. | Aplicación de la distancia de Gromov-Hausdorff al transporte público en AMBA | 87 |
| 6.7. | Casi distancia de Gromov-Hausdorff entre espacios casi-métricos | 92 |
| Capítulo 7. | Distancias entre medidas y distribuciones de probabilidad en espacios métricos | 93 |
| 7.1. | Introducción | 93 |
| 7.2. | Las métricas de Kantorovich-Rubinstein y de Wasserstein | 93 |

| | | |
|--------------|--|-----|
| 7.3. | La distancia variación total | 96 |
| 7.4. | Distancia de Hellinger | 97 |
| 7.5. | La protométrica de Bhattacharyya | 99 |
| 7.6. | La protométrica de Kullback-Leibler | 99 |
| 7.7. | Metrización de Kantorovich-Rubinstein de una clase de medidas probabilísticas discretas | 101 |
| 7.8. | Cálculo de algunas distancias entre distintos atributos en AMBA | 102 |
| | | |
| Capítulo 8. | Una casi-métrica de tipo Gromov-Lipschitz en clases de espacios métricos con medidas probabilísticas. | |
| | El espacio $(\mathcal{M}^l(X_0, d_0), \mathfrak{d}_{GL}^{KR})$ | 105 |
| 8.1. | Introducción | 105 |
| 8.2. | El espacio casi-métrico $(\mathcal{M}^l(X_0, d_0), \mathfrak{d}_{GL}^{KR})$ | 106 |
| 8.3. | Propiedades de continuidad de grafos métricos con atributos con la distancia de Gromov-Lipschitz-Kantorovich | 111 |
| 8.4. | Aplicación del enfoque Lipschitz-Gromov al análisis de SUBE en AMBA | 114 |
| | | |
| Capítulo 9. | El enfoque Hausdorff-Gromov en la proto-metrización de espacios métricos de probabilidad | 121 |
| 9.1. | Introducción | 121 |
| 9.2. | La proto-métrica de Gromov-Hausdorff-Kantorovich | 121 |
| 9.3. | Aplicación del enfoque de Hausdorff-Gromov a SUBE en AMBA | 125 |
| | | |
| Capítulo 10. | Distancia entre grafos de afinidades con atributos. Estabilidad del espectro del Laplaciano | 129 |
| 10.1. | Introducción | 129 |
| 10.2. | Distancia entre grafos con afinidades y atributos probabilísticos | 129 |
| 10.3. | Estabilidad del espectro del Laplaciano | 130 |
| 10.4. | Un ejemplo, perturbaciones del grafo bipartito | 133 |
| 10.5. | Aplicación al cálculo de las distancias entre las afinidades dadas por el transporte público en AMBA antes y durante ASPO-2020 por COVID-19 | 135 |
| | | |
| Capítulo 11. | Variables aleatorias valuadas en espacios métricos con medidas | 141 |
| 11.1. | Introducción | 141 |

| | | |
|--|--|-----|
| 11.2. | Variables aleatorias valuadas en $(\mathcal{M}^l(X_0, d_0), \mathfrak{d}_{GL}^{KR})$ | 141 |
| 11.3. | Grafos métricos aleatorios en un espacio métrico con atributos | 142 |
| 11.4. | Variables aleatorias valuadas en grafos de afinidades con atributos | 145 |
| 11.5. | Aleatorización de grafos clásicos | 146 |
| 11.6. | Aleatorización de afinidad y atributos en el grafo de SUBE | 149 |
| Capítulo 12. | Concentración de variables aleatorias valuadas en grafos métricos y de afinidades con atributos. Concentración espectral | 153 |
| 12.1. | Introducción | 153 |
| 12.2. | El método de Cramér-Chernoff. Variables aleatorias subgaussianas. Concentración | 153 |
| 12.3. | Variables aleatorias subgaussianas | 157 |
| 12.4. | Concentración de grafos con métricas y atributos aleatorios alrededor del grafo medio | 161 |
| 12.5. | Concentración espectral de grafos de afinidades con atributos aleatorios | 166 |
| Conclusiones | | 173 |
| Conclusiones teóricas centrales | | 173 |
| Conclusiones para la aplicación a SUBE en AMBA | | 173 |
| Bibliografía | | 175 |

Resumen

La motivación de los problemas centrales abordados es la Tesis tiene origen en la incursión del Grupo LABRA del IMAL en el análisis de la relación entre el Transporte Público en AMBA (Área Metropolitana Buenos Aires) y la propagación de COVID-19. Entre los numerosos problemas que ese análisis planteó y plantea, algunos se instalan directamente en lo que son las líneas centrales de aportes históricos del Grupo: Geometría de Espacios Métricos y Casi-métricos, Espacios de Tipo Homogéneo y Análisis Armónico Generalizado. Y también con temas en los que el grupo incursionó más recientes relacionados con probabilidades, como lo es el análisis de concentración y grandes desviaciones. En ese marco amplio, la construcción desde los datos de SUBE provistos por Nación Servicios, las matrices de conectividad entre los 41 distritos de AMBA se construyen haciendo un muestreo en ese conjunto grande de datos. Estas matrices, junto con atributos en cada distrito, dados por ejemplo por la cantidad de infectados en un momento determinado de la pandemia, definen un grafo no dirigido ponderado en sus aristas y con atributos en sus vértices. La aplicación de resultados de la Teoría Analítica de Grafos, como la Teoría Espectral del Laplaciano del contexto y las métricas difusivas inducidas, dependen entonces de esta selección aleatoria de los datos empíricos. Las preguntas teóricas generadas por este modelo son entonces el análisis de la concentración, las grandes desviaciones de los grafos aleatorios con respecto a los grafos medios y de sus correspondientes análisis espectrales. Esto plantea el problema de contar con definiciones adecuadas de distancias o métricas en familias de espacios métricos con medida, como lo son algunos grafos no dirigidos y ponderados en las aristas con atributos en los vértices. Los insumos de esta teoría son los de las metrificaciones de Gromov de espacios métricos y las de Kantorovich-Rubinstein-Wasserstein de espacios de medidas. Los insumos básicos de Teoría de Probabilidad son el método de Cramér-Chernoff y el Lema de Hoeffding. En el desarrollo de la Tesis algunas breves bifurcaciones del camino central produjeron resultados topológicos simples que pueden ser de utilidad en la generación de marcadores

empíricos en conjuntos de datos cuando no hay métricas naturales en el contexto. Los resultados más relevantes se resumen y aplican en el Capítulo 12 al que casi todos los demás capítulos apuntan y aportan. Es conveniente destacar aquí que se ha buscado ilustrar los conceptos y resultados teóricos a través de la implementación de algoritmos en el lenguaje Python y en particular considerando los datos de SUBE. Así, los principales resultados de esta Tesis asumen dos características bien diferenciadas y cooperativas. Por un lado, se introdujeron conceptos teóricos y se probaron resultados estructurales que los relacionan. Por otra parte se generaron algoritmos capaces de producir indicadores o marcadores que puedan caracterizar ciertas propiedades de conjuntos grandes de datos y en particular en lo relativo al transporte en AMBA a través de SUBE.

Introducción de la Tesis

Un grafo no dirigido, ponderado en las aristas y con atributos en los vértices es una 4-upla $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \bar{a}, W)$ donde $\mathcal{V} = \{1, \dots, n\}$ son los vértices, $\mathcal{E} = \{\{i, j\} : i \neq j \in \mathcal{V}\}$ son las aristas, $\bar{a} = (a_1, \dots, a_n)$ son los pesos en los vértices con $\sum_{i=1}^n a_i = 1$ y $a_i > 0$ para todo $i \in \mathcal{V}$ y $W = (w_{ij})_{i,j=1,\dots,n}$, $w_{ij} = w_{ji}$, $w_{ii} = 0$ y $w_{ij} \geq 0$ son ponderaciones de las aristas que pueden representar una métrica o una afinidad entre los vértices i y j . La teoría analítica espectral de grafos produce en \mathcal{G} un análisis armónico dado por su Laplaciano. La solución de la ecuación del calor en ese contexto produce difusiones que generan una manera natural de definir distancias entre vértices en \mathcal{V} . Todas estas magnitudes analíticas dependen de la elección inicial de atributos \bar{a} de los vértices y ponderaciones W de las aristas. Por otra parte la elección inicial suele ser intrínsecamente aleatoria. Por consiguiente, en vez de un grafo \mathcal{G} tenemos variables aleatorias valuadas en grafos, \mathcal{G}_ω . El objetivo central de esta tesis es estudiar la convergencia a cero cuando t tiende a infinito de las probabilidades de “lejanía” entre \mathcal{G}_ω y $\mathbb{E}(\mathcal{G})$ definido por las medias de $\bar{a}(\omega)$ y $W(\omega)$, precisamente

$$\mathcal{P}(\{\omega \in \Omega : \mathbf{d}(\mathcal{G}_\omega, \mathbb{E}(\mathcal{G})) > t\}).$$

Resulta claro que una buena definición de distancia entre grafos ponderados con atributos se hace necesaria. Esto plantea una buena parte del trabajo de esta Tesis. Pero también es de interés estudiar la convergencia a cero cuando t tiende a infinito, y su velocidad, de la probabilidad de que los operadores de Laplace de \mathcal{G}_ω y de $\mathbb{E}(\mathcal{G})$ difieran en más de t .

La Tesis se organiza con doce capítulos. Algunos de ellos son introductorios, generales y no pretenden originalidad. Hemos hecho un especial esfuerzo en ilustrar usando “NetworkX” de Python, algunas veces, generando nuestros propios algoritmos, casi todo los conceptos que introducimos y exploramos. Describimos a continuación la organización de la monografía.

En el Capítulo 1 nos proponemos introducir algunos aspectos de la teoría de grafos, el operador Laplaciano y su espectro. En la Sección 1.2 obtenemos el Laplaciano como el operador de Euler-Lagrange de una funcional de energía. En la Sección 1.3 introducimos el espectro del operador Laplaciano en grafos. En la Sección 1.4 introducimos varios ejemplos y en la 1.5 una de las matrices del transporte en AMBA que ilustrará varios de los capítulos posteriores.

Nuestra exposición asumirá, en el Capítulo 2, la forma clásica del desarrollo matemático de ideas cuantitativas asociadas a los problemas de metrización. En particular, aunque muchos de los resultados son conocidos, tomaremos como punto de partida la construcción de topologías a partir de afinidades. Es decir la construcción de topologías no mediadas por métricas. En este enfoque que toma *ab initio* la afinidad, la construcción de topologías a través del sistema de entornos juega un papel esencial, por ello es que en la sección siguiente del capítulo repasamos la idea de construcción de topología a través de sistemas de entornos.

En el Capítulo 3 nos proponemos estudiar la evolución del cardinal del conjunto de componentes conexas de un grafo no dirigido y ponderado en función de niveles de umbral en la afinidad que determina la relación entre vértices. Usaremos algoritmos en Python para contar las componentes conexas obtenidas y para obtener gráficas del número de componentes conexas como funciones del nivel de umbral. Aplicaremos esta estrategia al análisis del grafo del transporte en AMBA introducido en el Capítulo 1 y a algunas variaciones del mismo que tienen en cuenta otras nociones de vecindad aparte de aquellas determinadas por el transporte público.

Si bien los resultados del Capítulo 3 muestran estrechas relaciones entre los conceptos de métrica y afinidad, resultados cuantitativos precisos sólo se tienen bajo condiciones adecuadas de transitividad de la afinidad y algún tipo de triangularidad en la métrica. Por consiguiente el concepto ya clásico de distancia de Hausdorff entre conjuntos de un espacio métrico induce naturalmente un concepto asociado de afinidad entre esos conjuntos, cuando la afinidad es transitiva. En el caso general, esta relación se pierde, pero lo mismo es posible definir una afinidad de tipo Hausdorff entre conjuntos a partir de una afinidad entre los puntos que lo integran. El propósito del Capítulo 4 es doble, por un lado introducir el concepto y las propiedades básicas de afinidad de Hausdorff y

por el otro, repasar las de la métrica de Hausdorff que serán importantes en el desarrollo posterior. También proveeremos algunos ejemplos ilustrativos y alguna aplicación de estos conceptos al grafo de SUBE.

El objetivo del Capítulo 5 es introducir y ejemplificar una de las métricas en la familia que usaremos luego de todos los espacios métricos, en la parte central de la tesis, para metrizar la clase de los espacios métricos con medidas; la métrica de Gromov-Lipschitz.

En el Capítulo 6 introducimos el enfoque de Gromov-Hausdorff. Gromov propone sumergir los dos espacios métricos en análisis en un continente métrico mayor y en éste usar la distancia de Hausdorff, que expusimos en el Capítulo 4. En particular será posible calcular distancias entre grafos métricos con distintos números de vértices con las definiciones que introducimos en el Capítulo 6.

En consonancia con la amplitud y universalidad de usos y aplicaciones de la teoría general de la medida, los abordajes a diversas ideas de definir y calcular distancias entre medidas y distribuciones, provienen de campos muy diversos de la matemática. Estas investigaciones están hoy amplificadas por las aplicaciones de sus ideas básicas en Teoría de la Información, Aprendizaje Automático, Ciencia de Datos, etcétera. Los aportes intradisciplinarios más relevantes provienen de la Estadística Matemática, de la Teoría de Probabilidad, de la Teoría de la Información, del Análisis Matemático y de la Teoría Geométrica de la Medida. El Capítulo 7 está dedicado a introducir la métrica de Kantorovich-Rubinstein entre probabilidades que será de utilidad en el desarrollo posterior de la tesis, su relación con la métrica de Wasserstein y a la descripción breve de otras nociones usuales de distancias entre medidas, muchas de las cuales no alcanzan el rango matemático de métricas, pero resultan útiles en aplicaciones.

El Capítulo 8, al que hemos titulado “Una casi-métrica de tipo Gromov-Lipschitz en clases de espacios métricos con medidas probabilísticas. El espacio $(\mathcal{M}^l(X_0, d_0), \mathfrak{d}_{GL}^{KR})$ ”, es central para el desarrollo posterior de los aspectos teóricos de la tesis.

Los espacios métricos con medidas constituyen las extensiones naturales de los espacios Euclídeos equipados con la medida de Lebesgue que son los ambientes propicios para desarrollar y aplicar las técnicas del análisis armónico. Una estructura clásica en este sentido es la de espacio de tipo homogéneo, espacios casi-métricos con medidas duplicantes. Casos importantes, por sus aplicaciones, de tales contextos son los grafos en los que se

introducen métricas y medidas y, atendiendo al carácter aleatorio que a veces asumen estas estructuras, es conveniente y aún relevante en aplicaciones, tener buenas nociones de distancias entre grafos que atiendan a la doble estructura que soportan: métrica y medida. Hemos introducido en los capítulos anteriores varias nociones de distancias entre espacios métricos y entre medidas probabilísticas.

Usando las ideas de metrización de Gromov extendemos esta construcción para espacios métricos probabilísticos generales y nos interesamos especialmente, contemplando las aplicaciones en capítulos siguientes, en grafos métricos no dirigidos y con atributos dados.

En el Capítulo 9 recuperamos el enfoque de Hausdorff-Gromov para producir una plausible proto-métrica en clases de tipo \mathcal{M} de espacios métricos con medidas probabilísticas. No avanzaremos en ningún aspecto teórico de la idea y nos limitaremos a calcular la magnitud propuesta en algunos ejemplos sencillos y en nuestro persistente análisis de SUBE en AMBA.

En los capítulos anteriores habremos introducido e ilustrado, con ejemplos y algunas aplicaciones, varias nociones de distancias entre espacios métricos con medidas probabilísticas. En el Capítulo 10 se introduce también alguna noción útil y simple de métrica entre grafos de afinidades que también tengan atributos probabilísticos en los vértices. Algunas de las razones para hacerlo son el análisis de estabilidad del espectro del Laplaciano y su uso en el Capítulo 11 para introducir variables aleatorias valuadas en grafos de afinidades con atributos.

En el Capítulo 8 dotamos a $\mathcal{M}^l(X_0, d_0)$ con una estructura de espacio casi-métrico primero y esto nos permite considerar objetos aleatorios en esta super-estructura. Es decir, variables aleatorias valuadas en espacios métricos probabilísticos. Formalizamos esta idea en la Capítulo 11 y luego iremos al caso particular, pero importante, de los grafos métricos con atributos en los que tanto las métricas, como los atributos son variables aleatorias. También consideramos grafos de afinidades aleatorios. Finalmente ilustraremos la aleatoriedad en el espectro de grafos clásicos y del grafo de nuestra aplicación a SUBE en AMBA.

En el Capítulo 12 nos proponemos hacer un análisis de las grandes desviaciones de las medias para los objetos aleatorios que quedan definidos por variables aleatorias con

valores en $\mathcal{M}^l(X_0, d_0)$ como lo son $d_{GL}^{KR}((X(\omega), d(\omega), \mu(\omega)), (\bar{X}, \bar{d}, \bar{\mu}))$ y en el caso de grafos métricos y de afinidades, los Laplacianos aleatorios y sus autovalores. Usaremos el caso subgaussiano de la teoría de concentración de Cramér-Chernoff. El capítulo está organizado en cinco secciones. En las Secciones 12.2 y 12.3 expondremos los aspectos básicos del método de Cramér-Chernoff en el caso subgaussiano. En la Sección 12.4 consideramos la concentración de $d_{GL}^{KR}(\mathcal{G}(\omega), \mathcal{G}_0)$ siendo \mathcal{G}_0 el grafo medio de los $\mathcal{G}(\omega)$. En la Sección 12.5 estudiamos concentración del Laplaciano y del espectro y también un caso particular de concentración de las métricas difusivas asociadas.

Grafos no dirigidos y ponderados

1.1. Introducción

La teoría de grafos tiene una larga historia en combinatoria, con aplicaciones a la teoría de redes y la informática, así como a la teoría de números, la química y la física matemática. La profusión bibliográfica hace innecesario y aún imposible pretender siquiera exponer un punto de vista general del estado de avance de la teoría. Nuestro enfoque será puramente analítico y guiado por la teoría clásica de potencial.

En este capítulo nos proponemos introducir algunos aspectos de la teoría de grafos, el operador Laplaciano y su espectro. En la Sección 1.2 obtenemos el Laplaciano como el operador de Euler-Lagrange de un funcional de energía. En la Sección 1.3 introducimos el espectro del operador Laplaciano en grafos. En la Sección 1.4 expondremos varios ejemplos y en la Sección 1.5 una de las matrices del transporte en AMBA que ilustrará varios de los capítulos posteriores.

Comenzamos definiendo la estructura básica de grafos ponderados no dirigidos. Decimos que $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \bar{a}, W)$ es un grafo ponderado no dirigido, si $\mathcal{V} = \{1, 2, \dots, n\}$ es

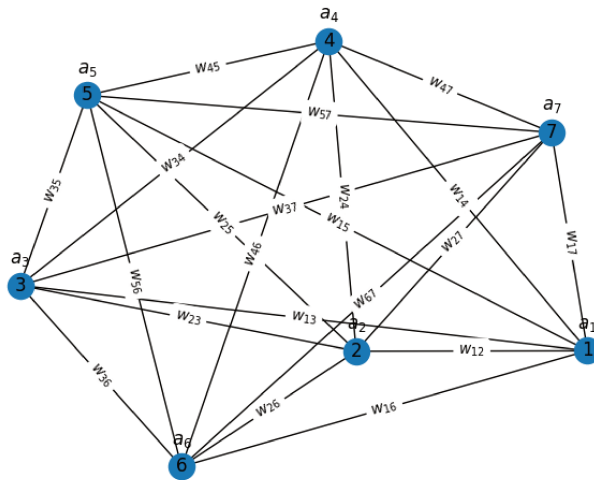


FIGURA 1.1.1. Un grafo ponderado no dirigido y con atributos.

un conjunto finito que llamamos nodos o vértices de \mathcal{G} , $\mathcal{E} = \{\{i, j\} : i, j \in \mathcal{V}\}$ es el conjunto de todos los subconjuntos de dos elementos de \mathcal{V} , llamados aristas de \mathcal{G} , \bar{a} es un vector de \mathbb{R}^n , $\bar{a} = (a_1, \dots, a_n)$ con $a_i > 0$ que es un valor de intensidad asignado de cada vértice i en \mathcal{V} , que llamamos atributo del vértice i . W es una matriz de $n \times n$ con componentes w_{ij} con $i, j = 1, \dots, n$ y $w_{ij} \geq 0$, que representa la afinidad o la distancia entre los nodos i y j del grafo \mathcal{G} . Cuando la matriz sea de afinidades y $w_{ij} = 0$ no habrá afinidad entre los nodos i y $j \in \mathcal{V}$. Como el carácter del grafo es no dirigido, la matriz W es una matriz simétrica, o sea $w_{ij} = w_{ji}$ para todo $i, j \in \mathcal{V}$. Consideramos que el grafo \mathcal{G} no tiene bucles, es decir que $w_{ii} = 0$, en otras palabras la diagonal de la matriz W es cero.

1.2. Energía de funciones reales definidas en los vértices de grafos no dirigidos ponderados. Ecuación de Euler-Lagrange

Recordemos que en un dominio Ω de \mathbb{R}^n , la energía de una función real definida en Ω está dada por la parte de regularidad de la norma de Sobolev de u , es decir

$$E(u) = \int_{\Omega} |\nabla u|^2 dx.$$

La minimización de esta forma cuadrática tiene como resultado la ecuación de Euler-Lagrange del contexto que produce el operador de Laplace clásico definido en coordenadas cartesianas como $\Delta u = \sum_{i=1}^n \frac{\partial^2 u}{\partial x_i^2}$.

En esta sección nos proponemos obtener el operador de Laplace en un grafo a partir de la minimización de una forma cuadrática de energía adecuada al contexto. Sea el grafo no dirigido ponderado $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \bar{a}, W)$ con $\#(\mathcal{V}) = n$ y supongamos que los nodos de \mathcal{V} tienen los mismos atributos y W es una matriz simétrica $n \times n$ con entradas no negativas y diagonal nula. Dada cualquier función con valores reales $f : \mathcal{V} \rightarrow \mathbb{R}$ y usando notación vectorial $f(i) = f_i$ definimos la energía de f con respecto a \mathcal{G} por

$$E(f) = \sum_{i=1}^n \sum_{j=1}^n w_{ij} (f_i - f_j)^2.$$

Esta noción de energía es la restricción a la diagonal $g = f$ de la forma bilineal

$$B(f, g) = \sum_{i=1}^n \sum_{j=1}^n w_{ij} (f_i - f_j)(g_i - g_j).$$

Para obtener la ecuación de Euler-Lagrange asociada a $E(f)$, como en el caso clásico consideramos la variación del funcional E . Esto se logra a través de la derivada de Gâteaux. Si $(V, \|\cdot\|)$ un espacio normado, U un abierto en V y $E : U \rightarrow \mathbb{R}$ una función. Dadas $f \in U$, $g \in V$, se dice que E es derivable Gâteaux en f en la dirección de g si el límite

$$\lim_{t \rightarrow 0} \frac{E(f + tg) - E(f)}{t}$$

existe y es finito. En ese caso el valor del límite se denota por $\delta E(f, g)$ y se llama la derivada (o variación) de Gâteaux de E en f en la dirección de g .

Tomando la derivada Gâteaux de la energía E en f en cada “dirección” g , en el siguiente resultado obtenemos el Laplaciano sobre \mathcal{G} como la ecuación de Euler-Lagrange asociada a E .

TEOREMA 1.2.1. *Sea $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \bar{a}, W)$ como antes. Si $f : \mathcal{V} \rightarrow \mathbb{R}$ es la que minimiza la energía E , entonces f es armónica con respecto al Laplaciano asociado al grafo \mathcal{G} , dado por*

$$\Delta_{\mathcal{G}} f(i) = \sum_{j=1}^n w_{ij} (f_i - f_j).$$

DEMOSTRACIÓN. Por definición de la derivada de Gâteaux de E en f en la dirección de g , tenemos

$$\delta E(f, g) = \lim_{t \rightarrow 0} \frac{1}{t} (E(f + tg) - E(f)).$$

La forma particular de E nos permite obtener una fórmula explícita para $\delta E(f, g)$. En efecto,

$$\begin{aligned} \frac{1}{t} (E(f + tg) - E(f)) &= \frac{1}{t} \left[\sum_{i=1}^n \sum_{j=1}^n w_{ij} (f_i + tg_i - f_j - tg_j)^2 - \sum_{i=1}^n \sum_{j=1}^n w_{ij} (f_i - f_j)^2 \right] \\ &= \frac{1}{t} \sum_{i=1}^n \sum_{j=1}^n [((f_i - f_j) + t(g_i - g_j))^2 - (f_i - f_j)^2] w_{ij} \\ &= \frac{1}{t} \sum_{i=1}^n \sum_{j=1}^n w_{ij} [2t(f_i - f_j)(g_i - g_j) + t^2(g_i - g_j)^2] \\ &= 2 \sum_{i=1}^n \sum_{j=1}^n (f_i - f_j) w_{ij} (g_i - g_j) + t \sum_{i=1}^n \sum_{j=1}^n (g_i - g_j)^2. \end{aligned}$$

Por lo tanto $\delta E(f, g) = 2 \sum_{i=1}^n \sum_{j=1}^n (f_i - f_j) \omega_{ij} (g_i - g_j)$. Como estamos suponiendo que f es un minimizador local de la energía E , debemos tener que $\delta E(f, g) = 0$ para cada g . Esta ecuación se traduce en

$$\sum_{i=1}^n \sum_{j=1}^n (f_i - f_j) \omega_{ij} (g_i - g_j) = 0$$

para cada $g : \mathcal{V} \rightarrow \mathbb{R}$. Entonces, para cada g ,

$$\begin{aligned} 0 &= \sum_{i=1}^n \sum_{j=1}^n \omega_{ij} (f_i - f_j) g_i - \sum_{i=1}^n \sum_{j=1}^n \omega_{ij} (f_i - f_j) g_j \\ &= \sum_{i=1}^n \sum_{j=1}^n \omega_{ij} (f_i - f_j) g_i + \sum_{i=1}^n \sum_{j=1}^n \omega_{ij} (f_j - f_i) g_j \\ &= \sum_{i=1}^n \sum_{j=1}^n \omega_{ij} (f_i - f_j) g_i + \sum_{i=1}^n \sum_{j=1}^n \omega_{ji} (f_i - f_j) g_i \\ &= 2 \sum_{i=1}^n \sum_{j=1}^n \omega_{ij} (f_i - f_j) g_i, \end{aligned}$$

en la última igualdad hemos usado la simetría $w_{ij} = w_{ji}$. De modo que

$$\sum_{i=1}^n \sum_{j=1}^n \omega_{ij} (f_i - f_j) g_i = 0$$

para cada g . Tomando $g : \mathcal{V} \rightarrow \mathbb{R}$ dada por $g(i) = g_i = 1$ y $g(l) = g_l = 0$ para $l \neq i$, obtenemos

$$\sum_{j=1}^n \omega_{ij} (f_i - f_j) = 0$$

□

Los mismos argumentos sirven para el caso mas general cuando los atributos no son todos iguales.

TEOREMA 1.2.2. *Sea $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \bar{a}, W)$ con \mathcal{V} , \mathcal{E} y W como antes. Supongamos ahora que cada vértice $i \in \mathcal{V}$ tiene un peso positivo a_i , por lo tanto $\bar{a} = (a_1, \dots, a_n)$. Si la energía E asociada a forma bilineal*

$$B(f, g) = \sum_{i=1}^n \sum_{j=1}^n \omega_{ij} (f_i - f_j) (g_i - g_j) a_i a_j$$

dada por

$$E(f) = B(f, f) = \sum_{i=1}^n \sum_{j=1}^n w_{ij} (f_i - f_j)^2 a_i a_j,$$

asume su mínimo en f , entonces

$$\sum_{j=1}^n w_{ij} (f_i - f_j) a_j = 0$$

para todo $i \in \mathcal{V}$.

DEMOSTRACIÓN. Ahora $E(f + tg) - E(f) = \sum_{i=1}^n \sum_{j=1}^n w_{ij} a_i a_j [2t(f_i - f_j)(g_i - g_j) + t^2(g_i - g_j)^2]$. Por lo tanto $\delta E(f, g) = 2 \sum_{i=1}^n \sum_{j=1}^n w_{ij} (f_i - f_j)(g_i - g_j) a_i a_j$. Como estamos suponiendo que f es un mínimo para E , entonces tenemos la ecuación

$$0 = \sum_{i=1}^n \sum_{j=1}^n w_{ij} a_i a_j (f_i - f_j)(g_i - g_j)$$

por cada $g : \mathcal{V} \rightarrow \mathbb{R}$. Como la matriz $(w_{ij} a_i a_j : i, j)$ es simétrica, tenemos que

$$\sum_{i=1}^n a_i \left[\sum_{j=1}^n w_{ij} (f_i - f_j) a_j \right] g_i = 0$$

para toda g . Entonces

$$\sum_{j=1}^n w_{ij} (f_i - f_j) a_j = 0$$

para todo $i \in \mathcal{V}$. □

1.3. El operador de Laplace en grafos ponderados no dirigidos con atributos

Los operadores “diferenciales” obtenidos en la sección anterior como minimizantes locales de un funcional de energía, corresponden en la teoría clásica de operadores elípticos a los operadores en forma de divergencia. Cuando, como en [BBL⁺17], pensamos los datos que la estructura del grafo nos provee en términos de espacios funcionales, y más precisamente de Hilbert, podemos ver a los datos $\bar{a} = (a_1, a_2, \dots, a_n)$ como una medida en \mathcal{V} y $W = (w_{ij} : i = 1, \dots, n; j = 1, \dots, n)$ como una medida en \mathcal{E} . Si bien, al ser ambos conjuntos \mathcal{V} y \mathcal{E} finitos, todos los espacios de funciones reales definidos sobre ellos coinciden, las medidas en \mathcal{V} y \mathcal{E} , definidas por \bar{a} y W respectivamente, inducen productos internos distintos de los usuales. Con respecto a estos pesos es conveniente conservar las propiedades de ortonormalidad que serán centrales en el desarrollo

de la teoría espectral. Sea $L^2(\mathcal{V}, \bar{a})$ el espacio vectorial de todas las funciones reales $f : \mathcal{V} \rightarrow \mathbb{R}$ equipado con el producto interno $\langle f, g \rangle_{\bar{a}} = \sum_{i=1}^n f_i g_i a_i$. De manera análoga $L^2(\mathcal{E}, W)$ es el espacio de todas las funciones reales $F : \mathcal{E} \rightarrow \mathbb{R}$ equipado con el producto interno $\langle F, G \rangle_W = \sum_{i=1}^n \sum_{j=1}^n F(\{i, j\}) G(\{i, j\}) w_{ij}$. En el espacio $L^2(\mathcal{V}, \bar{a})$ está bien definido un gradiente razonable como un operador que aplica $L^2(\mathcal{V}, \bar{a})$ en $L^2(\mathcal{E}, W)$. En efecto, dada $f \in L^2(\mathcal{V}, \bar{a})$ definimos ∇f como una función en $L^2(\mathcal{E}, W)$ dada por $\nabla f_{ij} = \nabla f(\{i, j\}) = f_j - f_i$. Por otra parte, si definimos otro operador D , ahora de $L^2(\mathcal{E}, W)$ en $L^2(\mathcal{V}, \bar{a})$ que a cada F de $L^2(\mathcal{E}, W)$ le asigna la función de $L^2(\mathcal{V}, \bar{a})$ dada por $(DF)_i = \frac{1}{a_i} \sum_{j=1}^n w_{ij} F(\{i, j\})$. Notemos que $\nabla f_{ij} = -\nabla f_{ji}$. Consideremos las funciones F de $L^2(\mathcal{E}, W)$ que cumplen $F_{ij} = -F_{ji}$.

PROPOSICIÓN 1.3.1. *El adjunto de ∇ es $2D$. Precisamente, para toda $F \in L^2(\mathcal{E}, W)$ y para toda $f \in L^2(\mathcal{V}, \bar{a})$ vale la identidad*

$$2\langle DF, f \rangle_{\bar{a}} = \langle F, \nabla f \rangle_W.$$

DEMOSTRACIÓN.

$$\begin{aligned} \langle DF, f \rangle_{\bar{a}} &= \sum_{i=1}^n (DF)_i f_i a_i \\ &= \sum_{i=1}^n \left(\frac{1}{a_i} \sum_{j=1}^n w_{ij} F(\{i, j\}) \right) a_i f_i \\ &= \sum_{i=1}^n f_i \sum_{j=1}^n w_{ij} F(\{i, j\}) \\ &= \sum_{i=1}^n f_i \sum_{j=1}^n w_{ij} F_{ij} \\ &= - \sum_{i=1}^n f_i \sum_{j=1}^n w_{ij} F_{ji} \\ &= - \sum_{j=1}^n f_j \sum_{i=1}^n w_{ji} F_{ij} \\ &= - \sum_{j=1}^n f_j \sum_{i=1}^n w_{ij} F_{ij}, \end{aligned}$$

restando la expresiones de la cuarta igualdad precedente y la séptima tenemos que

$$2\langle DF, f \rangle_{\bar{a}} = \sum_{j=1}^n \sum_{i=1}^n (f_j - f_i) w_{ij} F_{ij} = \langle \nabla f, F \rangle_W.$$

□

Ahora que tenemos un gradiente ∇ y un candidato a divergencia D , podemos definir un operador de Laplace de tipo Div Grad en campos escalares f definidos en \mathcal{V} por

$$(\Delta f)_i = (D(\nabla f))_i = \frac{1}{a_i} \sum_{j=1}^n w_{ij} (\nabla f)_{ij} = \frac{1}{a_i} \sum_{j=1}^n w_{ij} (f_j - f_i).$$

Sea $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \bar{a}, W)$ un grafo ponderado no dirigido como definimos en la introducción. La siguiente proposición describe en forma matricial el Laplaciano $\Delta_{\mathcal{G}}$ que será de utilidad en los cálculos y en particular en los ejemplos de la Sección 1.4.

PROPOSICIÓN 1.3.2. *Sea*

$$A = \begin{pmatrix} a_1 & \dots & 0 \\ & \ddots & \\ 0 & \dots & a_n \end{pmatrix}$$

con $a_i > 0$ para todo $i = 1, \dots, n$, las componentes del vector \bar{a} . Sea

$$W = \begin{pmatrix} 0 & w_{12} & w_{13} & \dots & w_{1n} \\ w_{12} & 0 & w_{23} & \dots & w_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{1n} & w_{2n} & w_{3n} & \dots & 0 \end{pmatrix}$$

con $w_{ii} = 0$ para todo $i = 1, \dots, n$ y $w_{ij} = w_{ji}$, la matriz de afinidades. Sea $f : \mathcal{V} \rightarrow \mathbb{R}$ dada por el n -vector de sus valores $f = (f_1, f_2, \dots, f_n)$. Sea

$$D = \begin{pmatrix} \sum_j w_{1j} & 0 & \dots & 0 \\ 0 & \sum_j w_{2j} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sum_j w_{nj} \end{pmatrix}.$$

Entonces

$$\Delta f = A^{-1}(W - D)f.$$

O , como operadores

$$(1.1) \quad \Delta = A^{-1}(W - D).$$

DEMOSTRACIÓN. Notemos que $(\Delta f)_i$ es la i -ésima componente del producto $A^{-1}(W - D)f$, en efecto.

$$\begin{aligned} \frac{1}{a_i} \sum_{j=1}^n (w_{ij} - D_{ij})f_j &= \frac{1}{a_i} \sum_{j=1}^n (w_{ij}f_j - D_{ij}f_j) \\ &= \frac{1}{a_i} \left(\sum_{j=1}^n w_{ij}f_j - \sum_{j=1}^n D_{ij}f_j \right) \\ &= \frac{1}{a_i} \left(\sum_{j=1}^n w_{ij}f_j - D_{ii}f_i \right) \\ &= \frac{1}{a_i} \left(\sum_{j=1}^n w_{ij}f_j - \left(\sum_{j=1}^n w_{ij} \right) f_i \right) \\ &= \frac{1}{a_i} \left(\sum_{j=1}^n w_{ij}(f_j - f_i) \right) \\ &= \Delta f(i). \end{aligned}$$

□

Denotamos con $\|\cdot\|_{\bar{a}}$ a la norma en $L^2(\mathcal{V}, \bar{a})$ inducida por el producto escalar $\langle \cdot, \cdot \rangle_{\bar{a}}$.

PROPOSICIÓN 1.3.3. *El operador Δ es autoadjunto con respecto al producto interno $\langle f, g \rangle_{\bar{a}}$.*

DEMOSTRACIÓN. Sean f y g dos funciones definida en \mathcal{V} , entonces

$$\begin{aligned} \langle \Delta f, g \rangle_{\bar{a}} &= \langle A^{-1}(W - D)f, g \rangle_{\bar{a}} \\ &= \sum_{i=1}^n (\Delta f)_i g_i a_i \\ &= \sum_{i=1}^n \left(\frac{1}{a_i} \sum_{j=1}^n (w_{ij} - D_{ij})f_j \right) g_i a_i \\ &= \sum_{i=1}^n \sum_{j=1}^n (w_{ij} - D_{ij})f_j g_i \end{aligned}$$

$$\begin{aligned}
&= \sum_{j=1}^n \sum_{i=1}^n (w_{ij} - D_{ij}) g_i f_j \\
&= \sum_{j=1}^n \left(\frac{1}{a_j} \sum_{i=1}^n (w_{ij} - D_{ij}) g_i \right) f_j a_j \\
&= \langle f, \Delta g \rangle_{\bar{a}}.
\end{aligned}$$

□

PROPOSICIÓN 1.3.4. *El operador $-\Delta$ es semidefinido positivo, en otras palabras $\langle -\Delta f, f \rangle_{\bar{a}} \geq 0$ para toda f .*

DEMOSTRACIÓN. Como antes, por la simetría de los w_{ij} , tenemos

$$\begin{aligned}
\langle -\Delta f, f \rangle_{\bar{a}} &= \sum_{i=1}^n (-\Delta f)_i f_i a_i \\
&= \sum_{i=1}^n \sum_{j=1}^n (D_{ij} - w_{ij}) f_j f_i \\
&= \sum_{i=1}^n D_{ii} f_i^2 - \sum_{i=1}^n \sum_{j=1}^n w_{ij} f_i f_j \\
&= \sum_{i=1}^n \sum_{j=1}^n w_{ij} f_i^2 - \sum_{i=1}^n \sum_{j=1}^n w_{ij} f_i f_j \\
&= \sum_{i=1}^n \sum_{j=1}^n w_{ij} (f_i^2 - f_i f_j) \\
&= \frac{1}{2} \left(\sum_{i=1}^n \sum_{j=1}^n w_{ij} (f_i^2 - f_i f_j) + \sum_{i=1}^n \sum_{j=1}^n w_{ij} (f_j^2 - f_i f_j) \right) \\
&= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n w_{ij} (f_i^2 + f_j^2 - 2f_i f_j) \\
&= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n w_{ij} (f_i - f_j)^2 \geq 0.
\end{aligned}$$

□

Retomando la ecuación (1.1) donde expresamos el Laplaciano en forma matricial como $\Delta_{\mathcal{G}} = A^{-1}(W - D)$ dado por la Proposición 1.3.2, el operador $\Delta_{\mathcal{G}}$ es una matriz $n \times n$ definida negativa y simétrica. De la teoría clásica del álgebra lineal sabemos que la matriz

$\Delta_{\mathcal{G}}$ tienen n autovalores λ reales, y es fácil probar que el primer autovalor es cero y los demás negativos.

COROLARIO 1.3.5. *El operador Δ es diagonalizable. Más precisamente, existe una sucesión de números no positivos λ_j , los autovalores y una sucesión ϕ_j , las autofunciones, de n vectores ortonormales con respecto al producto interno $\langle f, g \rangle_{\bar{a}}$ tal que los autovalores λ_j se pueden ordenar de tal manera que*

$$\lambda_{n-1} \leq \lambda_{n-2} \leq \dots \leq \lambda_1 < \lambda_0 = 0.$$

Además $\langle \phi_j, \phi_i \rangle_{\bar{a}} = \delta_{ij}$ y $\phi_0(k) = \frac{1}{\sum_{i=1}^n a_i} = 1$ con $k = 1, \dots, n$ y se satisfacen las ecuaciones espectrales

$$\Delta \phi_j = \lambda_j \phi_j, \quad j = 0, 1, \dots, n-1.$$

1.4. Ejemplos de grafos clásicos y sus espectros

En esta sección ilustramos ejemplos clásicos en la teoría de grafos, calculando el Laplaciano y su espectro.

EJEMPLO 1.4.1. Se dice que un grafo es bipartito si $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$ con $\mathcal{V}_1 \cap \mathcal{V}_2 = \emptyset$, es decir un grafo cuyos vértices se pueden separar en dos conjuntos disjuntos, de manera que las aristas no pueden relacionar vértices de un mismo conjunto, es decir no hay aristas con los dos vértices en \mathcal{V}_i para $i = 1, 2$.

Vértices = 9 - Aristas = 18

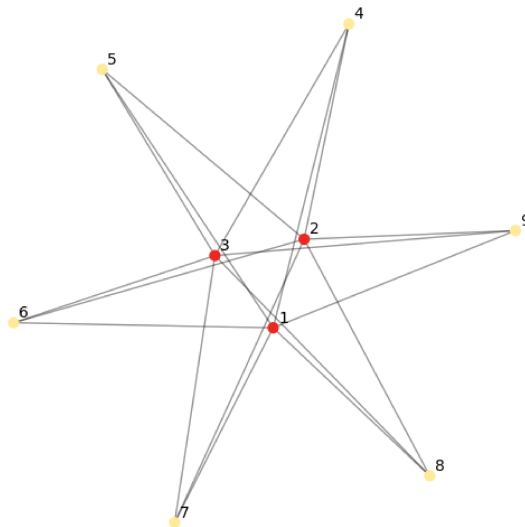


FIGURA 1.4.2. Grafo bipartito con $\#(\mathcal{V}_1) = 3$ y $\#(\mathcal{V}_2) = 6$.

Con la notación usada en la Sección 1.1 tenemos $\mathcal{V}_1 = \{1, 2, 3\}$, $\mathcal{V}_2 = \{4, 5, 6, 7, 8, 9\}$, $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$, $\mathcal{V}_1 \cap \mathcal{V}_2 = \emptyset$ con $\#(\mathcal{V}) = 9$. Sea $a_i = 1$, $i = 1, \dots, 9$; $w_{ij} = 1$ si $i \in \mathcal{V}_1$ y $j \in \mathcal{V}_2$ ó si $i \in \mathcal{V}_2$ y $j \in \mathcal{V}_1$ y $w_{ij} = 0$ si $i, j \in \mathcal{V}_1$ o $i, j \in \mathcal{V}_2$. Las matrices que definen el Laplaciano son las siguientes

$$A^{-1} = \bar{a} \cdot I_{9 \times 9} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad W = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Por razones que quedarán claras en el desarrollo posterior de la tesis, conviene normalizar también la matriz W a probabilidades. Es decir $\sum_{i=1}^n a_i = 1$ y $\sum_{j=1}^n \sum_{i=1}^n w_{ij} = 1$. En nuestro ejemplo, tomamos $\bar{a} = (\frac{1}{9}, \dots, \frac{1}{9})$ y $w_{ij} = \frac{1}{36}$ si $(i, j) \in (\mathcal{V}_1 \times \mathcal{V}_2) \cup (\mathcal{V}_2 \times \mathcal{V}_1)$.

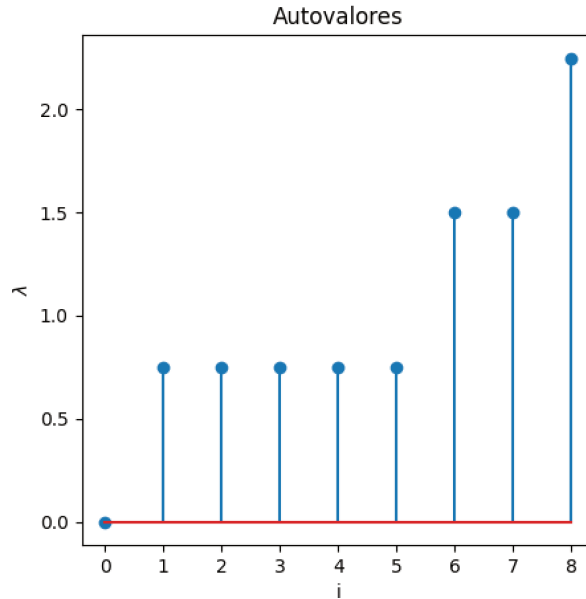
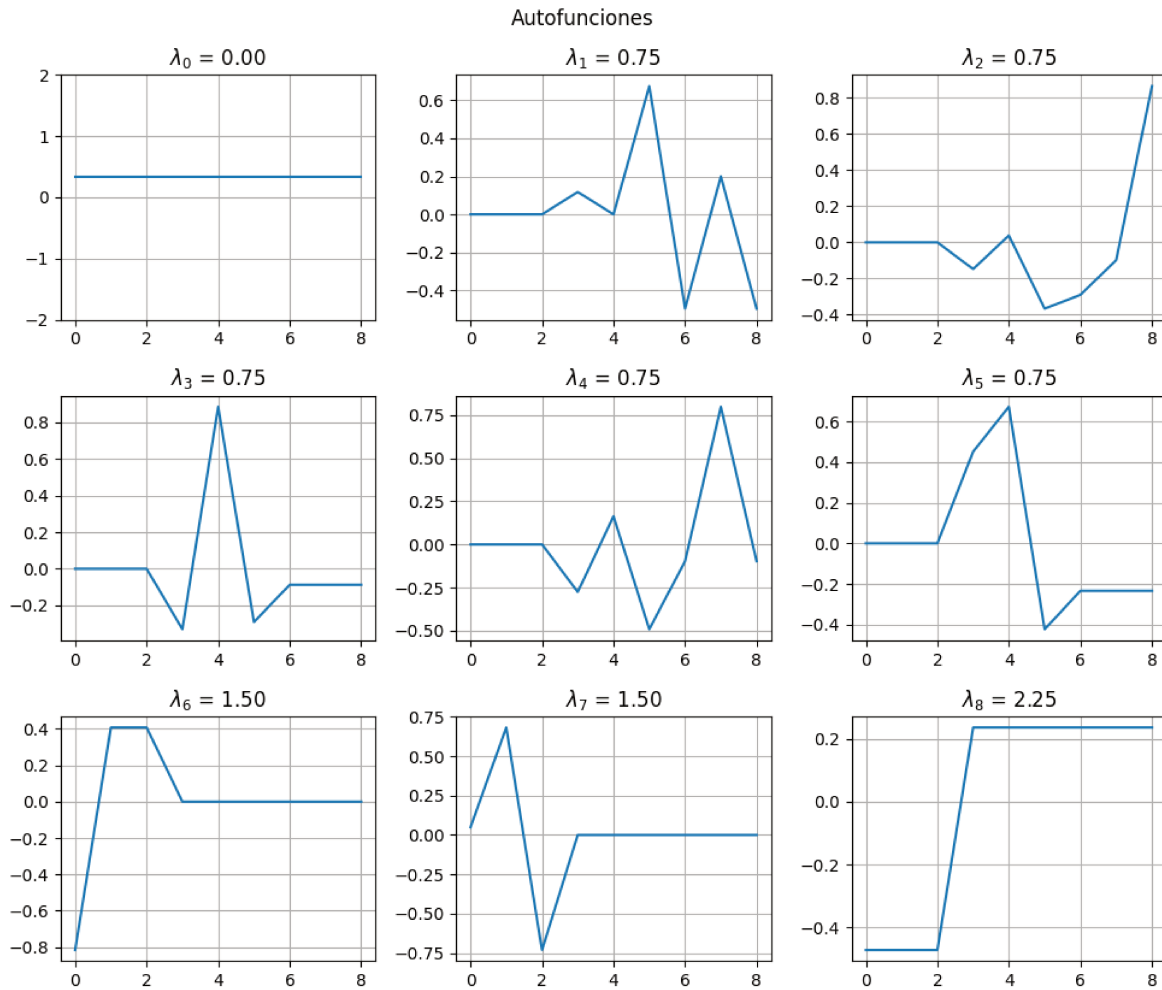
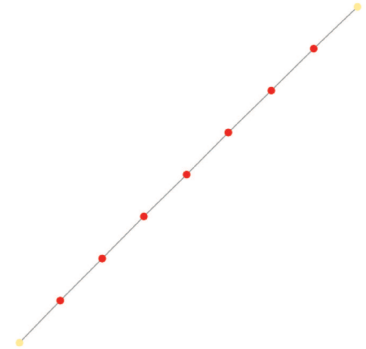
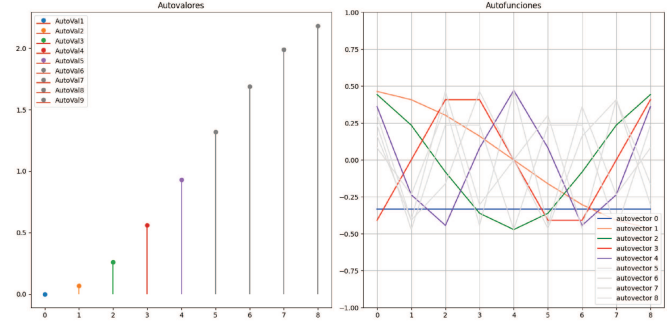
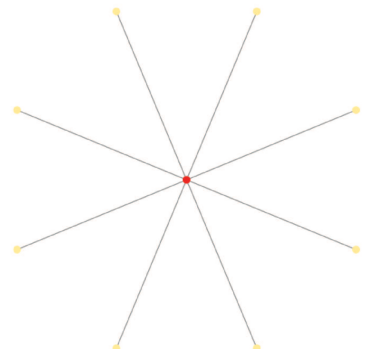
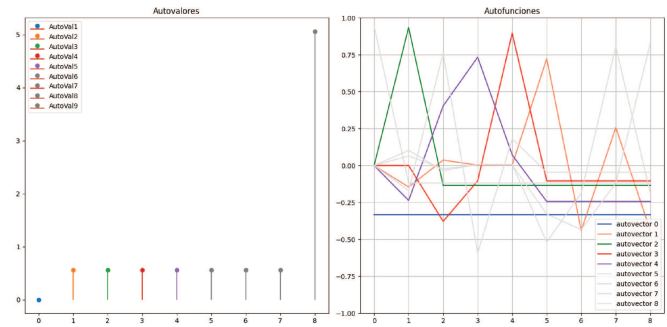
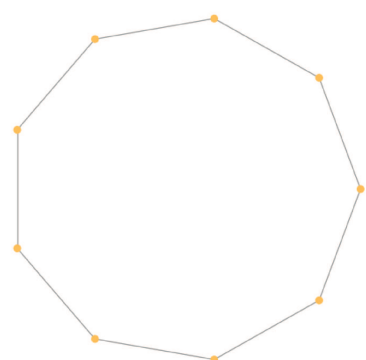
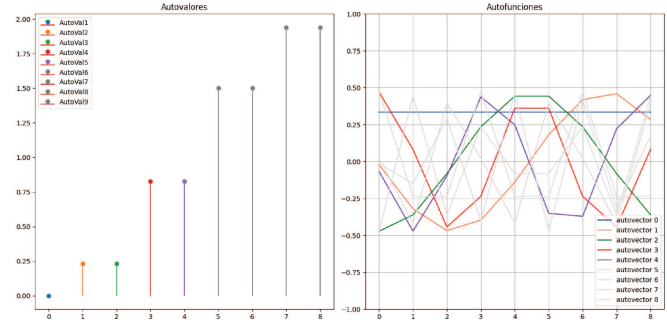


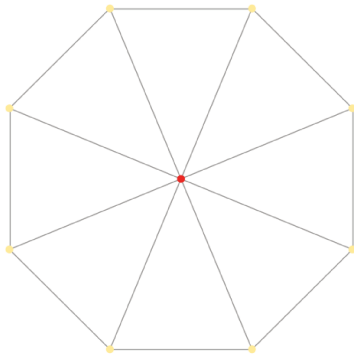
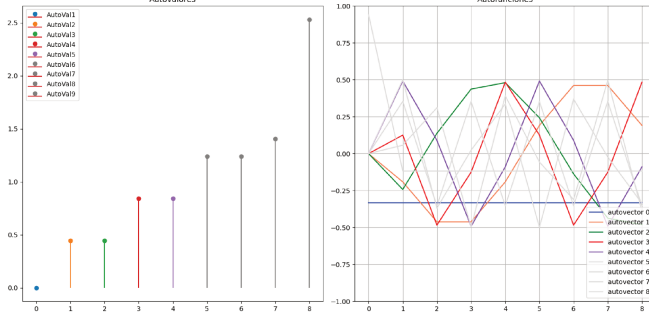
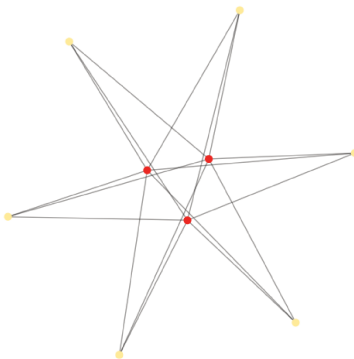
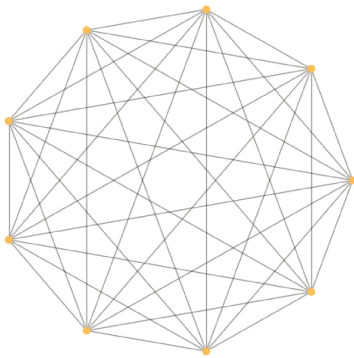
FIGURA 1.4.3. Autovalores de $-\Delta$.

FIGURA 1.4.4. Autofunciones de $-\Delta$.

Resumimos en la siguiente tabla las propiedades del espectro de algunos grafos clásicos, en cada caso se ilustra en colores diferenciados, las cuatro primeras autofunciones correspondientes a los cuatro primeros autovalores. Como en el caso del grafo bipartito utilizamos el lenguaje de Python para los cálculos del Laplaciano, de su espectro y la ilustración del mismo.

| Grafo | Esquema | Espectro |
|--------|---|--|
| Camino | <p>Vértices = 9 - Aristas = 8</p>  |  |
| | <p>Vértices = 9 - Aristas = 8</p>  |  |
| Ciclo | <p>Vértices = 9 - Aristas = 9</p>  |  |

CUADRO 1. Autovalores y Autofunciones de $-\Delta$.

| Grafo | Esquema | Espectro |
|----------|--|--|
| Rueda | <p>Vértices = 9 - Aristas = 16</p>  |  |
| | Bipartito | <p>Vértices = 9 - Aristas = 18</p>  |
| Completo | | <p>Vértices = 9 - Aristas = 36</p>  |

CUADRO 2. Autovalores y Autofunciones de $-\Delta$.

1.5. Un grafo para el transporte público en AMBA

Aplicamos los resultados de las secciones anteriores al análisis del transporte público de personas entre las 41 ciudades del AMBA en el contexto del COVID-19, a través de diferentes formas de transporte de pasajeros con el Sistema Único de Boleto Electrónico (SUBE). El acrónimo AMBA se utiliza para nombrar las 41 ciudades que contiene un tercio de la población total de Argentina y se concentran especialmente alrededor de la Ciudad de Buenos Aires. La Figura 1.5.5 muestra su localización geográfica y la etiqueta que le corresponderá en una numeración arbitraria pero fija a lo largo de la Tesis.

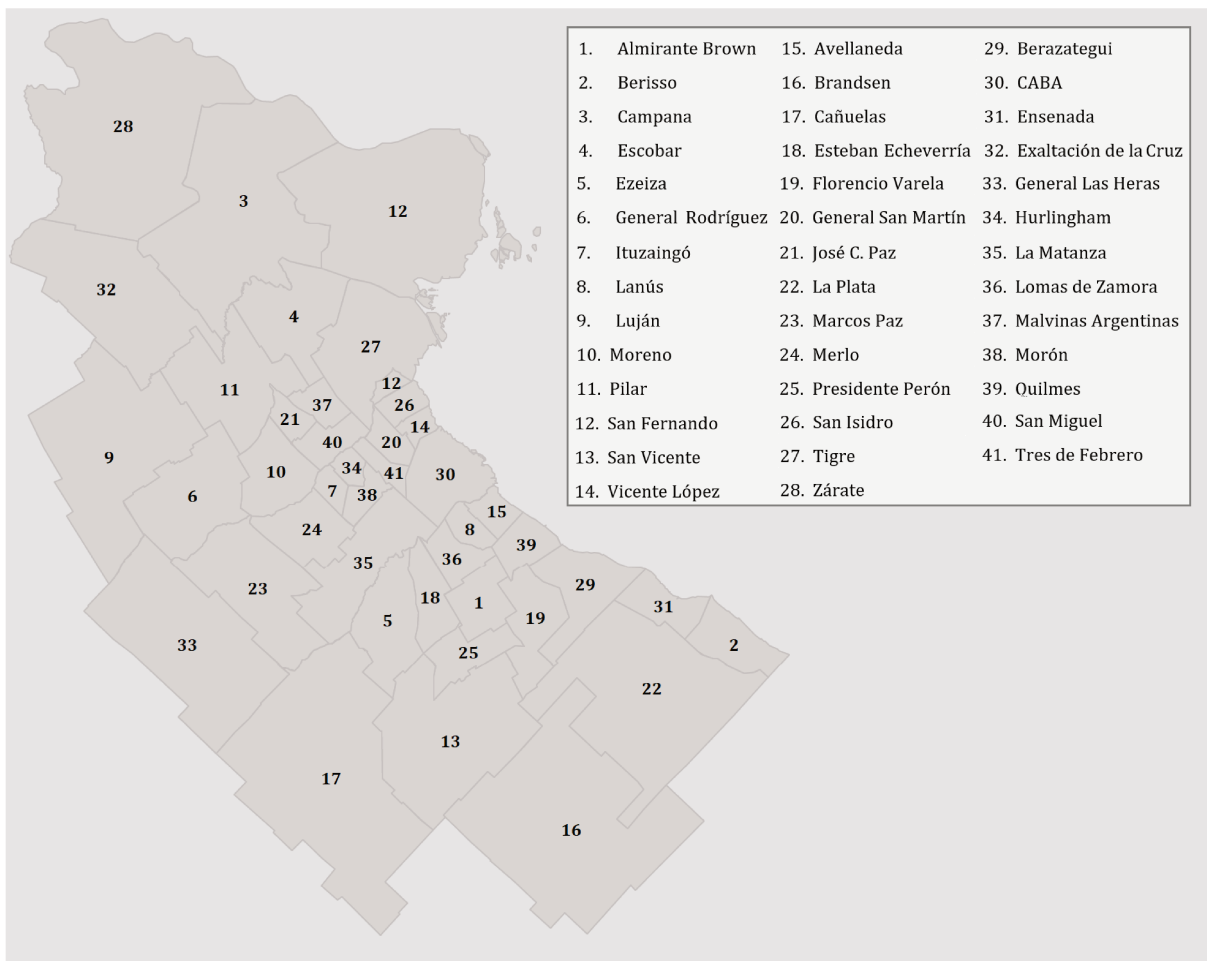


FIGURA 1.5.5. Mapa AMBA.

El sistema SUBE en AMBA integra los servicios de transporte de colectivos, metros, trenes y subterráneos del Área Metropolitana de Buenos Aires. Durante los primeros meses de evolución de la pandemia de COVID-19 en la Argentina, el grupo LABRA solicitó a Nación Servicios datos de SUBE en AMBA en etapas diferenciadas, previas al inicio de la pandemia y en vigencia de ASPO (Aislamiento Social Preventivo y Obligatorio). Los datos provistos fueron presentados en la forma “origen-destino” y cantidad de pasajeros que muestra un breve recorte de la planilla en la Figura 1.5.6.

| gid | Origen | gid | Destino | Viajes |
|-----|---------------|-----|--------------------|--------|
| 9 | Vicente L?pez | 9 | Vicente L?pez | 1298 |
| 9 | Vicente L?pez | 11 | Tigre | 235 |
| 9 | Vicente L?pez | 50 | General San Mart?n | 392 |
| 9 | Vicente L?pez | 52 | Ituzaing? | 6 |
| 9 | Vicente L?pez | 57 | San Isidro | 365 |
| 9 | Vicente L?pez | 62 | San Fernando | 70 |
| 9 | Vicente L?pez | 74 | Comuna 9 | 85 |
| 9 | Vicente L?pez | 97 | Avellaneda | 3 |
| 9 | Vicente L?pez | 101 | Comuna 15 | 119 |
| 9 | Vicente L?pez | 122 | Berazategui | 2 |
| 9 | Vicente L?pez | 145 | Comuna 7 | 27 |
| 9 | Vicente L?pez | 146 | Comuna 10 | 31 |
| 9 | Vicente L?pez | 147 | Comuna 11 | 55 |
| 9 | Vicente L?pez | 179 | Comuna 14 | 215 |
| 9 | Vicente L?pez | 192 | San Miguel | 56 |
| 9 | Vicente L?pez | 195 | Mercedes | 1 |
| 9 | Vicente L?pez | 197 | Marcos Paz | 2 |
| 9 | Vicente L?pez | 199 | Ezeiza | 3 |
| 9 | Vicente L?pez | 200 | Florencio Varela | 4 |
| 9 | Vicente L?pez | 260 | Escobar | 36 |
| 9 | Vicente L?pez | 299 | Comuna 13 | 373 |
| 9 | Vicente L?pez | 306 | Tres de Febrero | 40 |
| 9 | Vicente L?pez | 350 | La Plata | 3 |
| 9 | Vicente L?pez | 351 | Jos? C. Paz | 51 |
| 9 | Vicente L?pez | 354 | Moreno | 53 |
| 9 | Vicente L?pez | 355 | La Matanza | 62 |
| 9 | Vicente L?pez | 375 | Comuna 6 | 24 |
| 9 | Vicente L?pez | 390 | Comuna 1 | 451 |
| 9 | Vicente L?pez | 391 | Comuna 4 | 29 |
| 9 | Vicente L?pez | 392 | Comuna 5 | 41 |
| 9 | Vicente L?pez | 393 | Comuna 12 | 278 |

FIGURA 1.5.6. Recorte de 2245 datos de SUBE.

En su aporte al trabajo [AAGM22], Federico Morana logró extraer con la colaboración del Centro de Cómputos del CCT-Santa Fe, (Ricardo Plank) de esa profusión de datos varias matrices de conectividad entre los 41 distritos de AMBA. En particular la Figura 1.5.7 muestra esta matriz M para un día previo a la aplicación de ASPO: el 4 de marzo de 2020.

Es tal vez interesante preguntarse a cuál de los grafos contenidos en las Cuadro 1 y 2 resulta asimilable esta distribución de autovalores. Aunque resulta claro que no es estrictamente igual a ninguna de ellas observamos, considerando una estrella de 41 vértices con centro en el vértice 30, que las distribuciones de autovalores tienen un comportamiento cualitativo similar.

En la Figura 1.5.9 graficamos los autovalores de un grafo estrella con 41 vértices (color verde) y los autovalores que genera el Laplaciano del ejemplo AMBA (color azul).

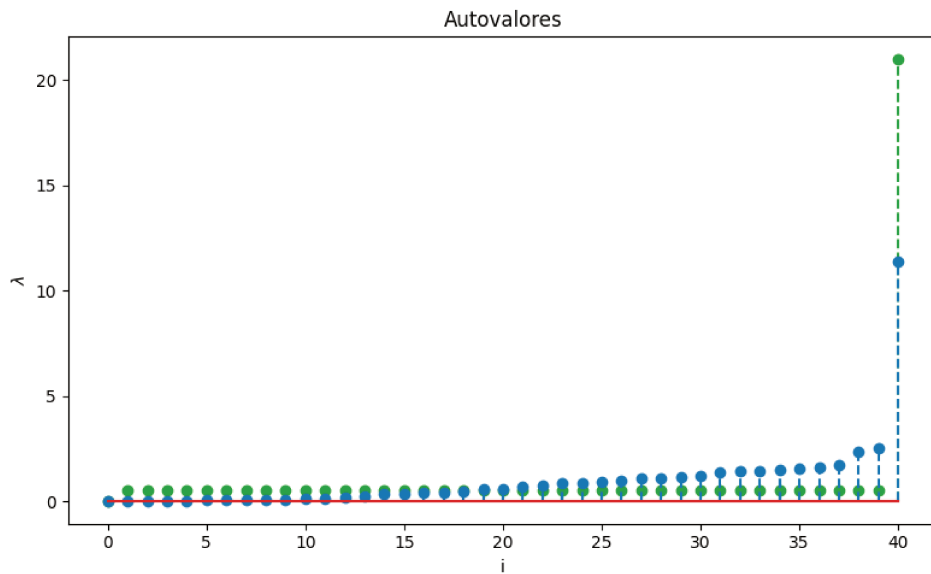


FIGURA 1.5.9. Comparación Espectral de SUBE en AMBA (04-03-20) con una estrella de 41 vértices y centro en el nodo 30.

Capítulo 2

Atracción, afinidad, topología y distancia

2.1. Introducción

El objeto de esta tesis es el estudio y uso de métricas y nociones de distancias entre objetos. Una situación que prefigura cierta circularidad, que rápidamente se atenúa y aún desaparece, es que nuestros objetos de interés que pretendemos detectar cuán cerca o lejos están, son a su vez espacios métricos. Para ser más precisos son espacios métricos con medida.

Por otra parte las tecnologías de la información y del análisis de datos han venido a llamar métrica a casi cualquier función que asigna números a pares de objetos. Sabemos que en matemática los requerimientos para llamar métricas a tales asignaciones son algo más exigentes. Esperamos de ellas que generen topologías, uniformidades, nociones de convergencia. La matemática ha tenido el problema de la metrización de espacios topológicos y uniformes como una de las cuestiones centrales durante la primera mitad del Siglo XX. Estos resultados cualitativos tienen correlatos cuantitativos que son relevantes en los problemas de metrización asociados a conjuntos de datos.

Nuestra exposición asumirá, en este capítulo, la forma clásica del desarrollo matemático de ideas cuantitativas asociadas a los problemas de metrización. En particular aunque muchos de los resultados son conocidos, tomaremos como punto de partida la construcción de topologías a partir de afinidades. Es decir la construcción de topologías no mediadas por métricas. En este enfoque que toma *ab initio* la afinidad, la construcción de topologías a través del sistema de entornos juega un papel esencial, por ello es que en la sección siguiente del capítulo repasamos la idea de construcción de topología a través de sistemas de entornos.

2.2. Topología y sistemas de entornos

Dado un conjunto X , una **topología** en X es una familia τ de las partes de X que es cerrada por uniones arbitrarias e intersecciones finitas y que contiene al vacío y al propio

conjunto X . Los elementos de τ se llaman **abiertos**. El par (X, τ) se llama **espacio topológico**. Para $x \in X$, un subconjunto V de X es un **entorno** de x si existe $A \in \tau$ tal que $x \in A \subset V$. Un subconjunto F de X es **cerrado** si $F^c \in \tau$. La intersección de todos los cerrados que contienen a E es la clausura de E que se denota \bar{E} . Un subconjunto D de X es **denso** si $\bar{D} = X$. El espacio topológico es **separable** si existe un subconjunto denso y numerable en X . Una familia \mathcal{B} de subconjuntos de X es una **base** para τ si $\mathcal{B} \subset \tau$ y cada A en τ es unión de elementos de \mathcal{B} . Los espacios topológicos con bases numerables son separables.

Si (X, τ) es un espacio topológico y sea \mathcal{N} la función que a cada $x \in X$ le asigna la familia \mathcal{N}_x de todos los entornos de x , entonces se tienen las siguientes propiedades básicas

- i) $U \in \mathcal{N}_x$ implica $x \in U$;
- ii) $U, V \in \mathcal{N}_x$ implica $U \cap V \in \mathcal{N}_x$;
- iii) $U \in \mathcal{N}_x$ y $V \supset U$ implica $V \in \mathcal{N}_x$.

El siguiente resultado puede verse en [Kel62]. Por razones de completitud agregamos una demostración.

PROPOSICIÓN 2.2.1. *Si X es un conjunto y \mathcal{N} una función que cumple i), ii) y iii), entonces $\tau = \{U : U \in \mathcal{N}_x \text{ para todo } x \in U\}$ es una topología en X .*

DEMOSTRACIÓN. Es claro que \emptyset y X pertenecen a τ . Asimismo es simple ver que intersecciones finitas y uniones arbitrarias están en τ . En efecto, sean $U_1, U_2 \in \tau$. Sea $z \in U_1 \cap U_2$, entonces U_1 y $U_2 \in \mathcal{N}_z$. Por consiguiente por ii) $U_1 \cap U_2 \in \mathcal{N}_z$ se tiene que $U_1 \cap U_2 \in \tau$. Consideremos ahora $\{U_\alpha, \alpha \in A\} \subset \tau$. Sea $x \in \bigcup_{\alpha \in A} U_\alpha$ entonces existe un α_0 tal que $x \in U_{\alpha_0} \in \tau$. Entonces $U_{\alpha_0} \in \mathcal{N}_x$ y como $U_{\alpha_0} \subset \bigcup_{\alpha \in A} U_\alpha$, por iii) tenemos $\bigcup_{\alpha \in A} U_\alpha \in \mathcal{N}_x$. De aquí que $\bigcup_{\alpha \in A} U_\alpha \in \tau$. □

Entre las propiedades de un espacio topológico y de sus subconjuntos, sólo abordaremos brevemente el repaso de la propiedad de conectividad que será explorada en el contexto de grafos ponderados. Se dice que (X, τ) es **conexo** si no se puede escribir como unión disjunta de dos abiertos no vacíos. Precisamente (X, τ) es **conexo** si no existen U y $V \in \tau$ con $U \neq \emptyset$, $V \neq \emptyset$, $U \cap V = \emptyset$ tales que $X = U \cup V$. Un subconjunto A en el espacio topológico (X, τ) es conexo si (A, τ_A) es un espacio topológico conexo con τ_A la topología

que A hereda de (X, τ) , $\tau_A = \{U \cap A : U \in \tau\}$. Un conjunto conexo maximal en A se llama **componente conexa** de A . Es claro que A es conexo si y sólo si $\#(\mathcal{C}(A, \tau_A)) = 1$, donde $\mathcal{C}(A, \tau_A)$ es el conjunto de todas las componentes conexas de (A, τ_A) .

PROPOSICIÓN 2.2.2. *Sea X un conjunto.*

- (a) *Si τ es una topología en X entonces $1 \leq \#(\mathcal{C}(X, \tau)) \leq \#(X)$.*
- (b) *Si $\tau_1 \subseteq \tau_2$ son dos topologías en X entonces $\#(\mathcal{C}(X, \tau_1)) \leq \#(\mathcal{C}(X, \tau_2))$.*

DEMOSTRACIÓN. La parte (a) se deduce del hecho que las topologías minimal y maximal son precisamente $\{\emptyset, X\}$ y $\mathcal{P}(X)$. Probemos (b). Notemos primero que $\tau_1 \subset \tau_2$ implica que todo τ_2 -conexo es τ_1 -conexo. Por consiguiente toda componente conexa en τ_2 debe estar contenida en alguna componente conexa en τ_1 . Por consiguiente $\#(\mathcal{C}(X, \tau_1)) \leq \#(\mathcal{C}(X, \tau_2))$. \square

2.3. Topologías inducidas por protométricas

El concepto de protométrica (ver [Aim]) es, de alguna manera, el más elemental y rudimentario posible para generar una topología en un conjunto a través de un sistema de entornos de cada punto. La Proposición 2.2.1 se aplica cuando se tiene una noción de distancia que no cumple las propiedades de una métrica.

Sea X un conjunto no vacío. Por simplicidad notacional, usaremos \mathbb{R}_0^+ para denotar el conjunto de los números reales no negativos. Una función $g : X \times X \rightarrow \mathbb{R}_0^+$ es una protométrica si $g \equiv 0$ en la diagonal Δ de $X \times X$. Es decir si $g(x, x) = 0$ para todo $x \in X$. Notar que no hay condiciones de simetría, de confiabilidad (anulación sólo sobre Δ), ni tampoco ningún requerimiento de tipo desigualdad triangular. Esta austeridad no impide el uso de g para generar una topología en X a través de un sistema de entornos de cada punto que, de alguna manera, esté controlada por la función g .

Sea $B_g(x, r) = \{y \in X : g(x, y) < r\}$ la (proto)bola en $x \in X$ con radio $r > 0$. Sea $\mathcal{N} : X \rightarrow \mathcal{P}(\mathcal{P}(X))$ ($\mathcal{P}(A)$ = partes de A) dado por $\mathcal{N}(x) = \mathcal{N}_x$, con $\mathcal{N}_x = \{U : \text{para algún } r > 0 \text{ vale que } B_g(x, r) \subset U\}$.

PROPOSICIÓN 2.3.1. *Sea g una protométrica en X y \mathcal{N} la función definida arriba. Entonces \mathcal{N} satisface i), ii) y iii) de sistemas de entornos y*

$\tau = \tau_g = \{U : \text{si } x \in U, \text{ existe } r > 0 \text{ tal que } B_g(x, r) \subset U\}$ es una topología en X .

DEMOSTRACIÓN. *i)* $U \in \mathcal{N}_x$ implica que para algún $r > 0$, $y \in U$ si $g(x, y) < r$. Pero $0 = g(x, x) < r$. Entonces $x \in U$. *ii)* U se tiene que $V \in \mathcal{N}_x$ implica que existen r_1 y r_2 positivos con $B_g(x, r_1) \subset U$ y $B_g(x, r_2) \subset V$. Entonces como $B_g(x, \min\{r_1, r_2\}) \subset U \cap V$ y $U \cap V \in \mathcal{N}_x$. *iii)* es clara. \square

Dos casos extremos de la situación anterior generada por proto-métricas en X dan también las topologías extremas en X . Si $g \equiv 0$, $\tau_g = \{\emptyset, X\}$ y si g es la función característica del complemento de la diagonal de $X \times X$ la topología τ_g es la más grande, $\tau_g = \mathcal{P}(X)$.

2.4. Topologías inducidas por afinidades en conjuntos arbitrarios

En esta sección usaremos el resultado de la Sección 2.2 para construir topologías cuando en el conjunto básico X , en vez de una métrica o sus variantes, tenemos una noción de afinidad.

Newton, en su teoría de gravitación, descubrió que la fuerza de atracción (afinidad) entre dos masas es inversamente proporcional al cuadrado de la distancia euclídea que las separa. En este sentido es que se puede pensar a métricas y afinidades en una relación recíproca.

Un ejemplo simple y básico de afinidad en espacios euclídeos, que puede verse en [CL06], sobre un conjunto finito de puntos $\{x_i : i = 1, \dots, N\}$ de \mathbb{R}^n , está dado por el perfil gaussiano $w_{ij}^\epsilon = e^{-\frac{\|x_i - x_j\|^2}{\epsilon}}$, $i, j = 1, \dots, N$ y $\epsilon > 0$ fijo. En este caso la máxima afinidad ($= 1$) ocurre sólo cuando $x_i = x_j$. En este ejemplo la máxima afinidad es uno y corresponde a la diagonal. Seremos más drásticos en nuestra definición.

En general, dado un conjunto X cualquiera diremos que una función $W : X \times X \rightarrow \mathbb{R}_*^+ = \mathbb{R}^+ \cup \{+\infty\}$ es una **afinidad** en X si $W(x, x) = +\infty$, para todo $x \in X$ donde \mathbb{R}^+ denota el conjunto de los números reales positivos. Usaremos a veces, en particular en los casos discretos, la notación $W_{xy} = W(x, y)$, $(x, y) \in X \times X$. Si W es una afinidad en X diremos que el par (X, W) es un **espacio con afinidad**. Todo espacio con afinidad resulta un espacio topológico. El resultado general de la Sección 2.2 nos permite probar el siguiente teorema.

TEOREMA 2.4.1. *Sea X un conjunto. Sea W una afinidad en X . Entonces la función $x \mapsto \mathcal{N}_x$, con $\mathcal{N}_x = \{U \subset X : \text{existe } \alpha > 0 \text{ tal que } E(x, \alpha) = \{y : W(x, y) > \alpha\} \subset U\}$*

satisface *i*), *ii*) y *iii*) de la Sección 2.2. Por consiguiente

$$\begin{aligned}\tau_W &= \{U \subset X : U \in \mathcal{N}_x \text{ para todo } x \in U\} \\ &= \{U \subset X : \text{para todo } x \in U \text{ existe } \alpha = \alpha(x) > 0 \text{ tal que } E(x, \alpha) \subset U\}\end{aligned}$$

es una topología en X .

DEMOSTRACIÓN. *i*) $U \in \mathcal{N}_x$ entonces para algún $\alpha > 0$ se verifica que $y \in U$ si $W(x, y) > \alpha$. Pero $W(x, x) = +\infty > \alpha$. Entonces $x \in U$. *ii*) U y $V \in \mathcal{N}_x$ implica que existen $\alpha_1 > 0$ y $\alpha_2 > 0$ tal que $E(x, \alpha_1) \subset U$ y $E(x, \alpha_2) \subset V$. Entonces $E(x, \max\{\alpha_1, \alpha_2\}) \subset U \cap V$ y por definición $U \cap V \in \mathcal{N}_x$. Para *iii*) tenemos que $U \in \mathcal{N}_x$ esto implica que existe un $\alpha > 0$ tal que $E(x, \alpha) \subset U$ y como $U \subset V$ entonces $E(x, \alpha) \subset V$ por definición de \mathcal{N}_x tenemos $V \in \mathcal{N}_x$. \square

2.5. Métricas, casi-métricas y pseudométricas

El caso más conocido y útil desde el punto de vista de la matemática de las protométricas introducidas en la Sección 2.3 es el de las métricas. Se dice que una función $d : X \times X \rightarrow \mathbb{R}_0^+$ es una **métrica o distancia** en X si satisface las siguientes propiedades

- $M_1)$ $d(x, y) \geq 0$ para todo x, y en X ;
- $M_2)$ $d(x, y) = 0$ si y sólo si $x = y$;
- $M_3)$ $d(x, y) = d(y, x)$ para todo x, y en X ;
- $M_4)$ $d(x, z) \leq d(x, y) + d(y, z)$ para todo x, y y z en X .

Diremos en cambio que d es una **casi-métrica** si d satisface las propiedades $M_1)$, $M_2)$ y $M_3)$ anteriores y en vez de la desigualdad triangular se satisface la siguiente forma débil de la misma

$$CM_4) \text{ existe alguna constante } K > 1 \text{ tal que } d(x, z) \leq K(d(x, y) + d(y, z)).$$

Por otra parte cuando d satisface solo $M_1)$, $M_3)$ y $M_4)$ se dice que es **pseudométrica** y cuando d satisface $M_1)$, $M_3)$ y $CM_4)$ es una **pseudo-casi-métrica**. Como todas estas nociones de distancia satisfacen $M_1)$, todas son protométricas por lo tanto todas inducen topologías en X de acuerdo a la Sección 2.3.

Los casos de pseudométrica o pseudo-casi-métrica merecen especial atención ya que en general producen en X topologías no metrizable. Por otra parte la inclusión del concepto de casi-métrica genera mucha mayor amplitud de modelos de análisis en ecuaciones

en derivadas parciales y en muchas otras aplicaciones. Sin embargo, y a pesar de su generalidad, un célebre teorema de Macías y Segovia del 1979, demuestra que las topologías generadas por casi-métricas son metrizable. Más que eso, las casi-métricas son equivalentes a potencias de métricas, es decir, sea ρ una casi-métrica en X con constante triangular mayor que uno, entonces existe una métrica d en X y un número $\beta \geq 1$ tal que

$$c_1\rho(x, y) \leq d^\beta(x, y) \leq c_2\rho(x, y)$$

para cada $x, y \in X$ y constantes $0 < c_1 \leq c_2 < \infty$ que dependen sólo de la constante triangular para ρ [MS79].

La demostración de este teorema se basa en el de metrización de espacios uniformes con bases numerables que puede encontrarse en el Capítulo 6 del libro de Kelley [Kel62] de topología general. A su vez este resultado en Kelley es una aplicación de un notable lema de Aline Huke, que vino a conocerse como el Lema de Frink por el apellido de casada de Aline.

Diremos que dos casi-métricas d, ρ son equivalentes cuantitativamente cuando existen dos constantes positivas c_1 y c_2 tales que para todo $x, y \in X$

$$c_1d(x, y) \leq \rho(x, y) \leq c_2d(x, y).$$

2.6. Metrización de afinidades transitivas

En la Sección 2.4 hemos visto que todo espacio con afinidad (X, W) tiene una topología natural asociada a W . Esta topología no siempre resulta metrizable ya que en general no cumple el axioma de separación de Hausdorff. El Lema de Aline Huke (Frink) también puede aplicarse a la metrización de espacios con afinidad especiales. Un resultado en esta dirección puede verse en [AG18]. Un punto de partida heurístico para esto es, como observamos antes, el de la relación entre la métrica euclídea clásica con la atracción gravitatoria considerada como afinidad. En efecto, si tomamos el punto de vista del potencial de Newton en vez de la fuerza, tenemos que la forma básica del mismo es $N(x, y) = \frac{1}{|x-y|}$, salvo constantes multiplicativas. Aquí x e y están en \mathbb{R}^3 y $|x - y|$ es su distancia euclídea. Notemos que $N > 0$ en todo $\mathbb{R}^3 \times \mathbb{R}^3$ y que $N = +\infty$ en la diagonal de $\mathbb{R}^3 \times \mathbb{R}^3$. Por consiguiente N es una afinidad en \mathbb{R}^3 en el sentido de la Sección 2.4 precedente. Pero por cierto N tiene otras propiedades, por ejemplo $N(x, y) = N(y, x)$. Y

también N cumple una propiedad transitiva que resulta natural pensando en afinidades y es la siguiente.

LEMA 2.6.1. *Sea $\lambda > 0$. Entonces para toda terna x, y, z de puntos del espacio \mathbb{R}^3 para los cuales $N(x, y) > \lambda$ y $N(y, z) > \lambda$, se tiene que $N(x, z) > \frac{1}{2}\lambda$.*

DEMOSTRACIÓN. $N(x, y) > \lambda$ implica $|x - y| < \frac{1}{\lambda}$ y $N(y, z) > \lambda$ implica $|y - z| < \frac{1}{\lambda}$. Por consiguiente $|x - z| \leq |x - y| + |y - z| < \frac{2}{\lambda}$ y entonces $N(x, z) = \frac{1}{|x-z|} > \frac{\lambda}{2}$. \square

En [AG18] se consideran afinidades que extienden estas propiedades básicas de N y que resultan suficientes para la metrización del espacio. Sea X un conjunto. Se dice que $W : X \times X \rightarrow \mathbb{R}_*^+$ es una **afinidad transitiva** si W es simétrica tal que satisface la siguiente propiedad: existe $\nu : \mathbb{R}^+ \rightarrow \mathbb{R}^+$, creciente, continua y cóncava que satisface $\nu(t) < t$ para todo $t > 0$ y tal que si $x, y, z \in X$ y $\lambda > 0$ satisfacen las desigualdades

$$W(x, y) > \lambda \quad \text{y} \quad W(y, z) > \lambda,$$

se tiene que

$$W(x, z) > \nu(\lambda).$$

En el ejemplo del potencial de Newton $\nu(t) = \frac{t}{2}$ es lineal. El resultado principal en [AG18] es el siguiente.

TEOREMA 2.6.2. *Sea X un conjunto y W una afinidad transitiva en X . Entonces existe una función decreciente y continua $\varphi : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ y una casi-métrica d en X tal que*

$$W(x, y) = \varphi(d(x, y)).$$

Cuando, como en el ejemplo del potencial de Newton, $\nu(t) = \frac{t}{2}$ es lineal, la función φ es una potencia. En general la construcción del perfil φ está contenido en la Proposición 3.1 de [AG18].

Bajo las condiciones del Teorema 2.6.2 si (X, W) es un espacio con afinidad transitiva, entonces la topología inducida por el Teorema 2.4.1 y la inducida por la casi-métrica d del Teorema 2.6.2 coinciden y entonces el espacio con afinidad (X, W) tendrá todas las propiedades topológicas y métricas del espacio (X, d) .

2.7. Métricas Difusivas

Introducimos en esta sección una familia de métricas inducidas por el análisis espectral de un grafo. Las métricas difusivas (Ver [CL06]). Si en una placa de material no homogéneo están dados dos puntos x e y , y en cada uno de ellos se introduce una unidad de energía térmica (calor), es decir δ_x y δ_y (deltas de Dirac), se deja pasar un tiempo $t > 0$, y luego se toma una termografía se obtienen distribuciones de temperatura $\theta_x^t(z)$ y $\theta_y^t(z)$, para z en la placa. Resulta que la cantidad $\sqrt{\int_{\text{Placa}} |\theta_y^t(z) - \theta_x^t(z)|^2 dz} = m(x, y)$ es una métrica en la placa. Esta métrica refleja mejor las eventuales inhomogeneidades y anisotropías del material que la compone y provee una distancia entre x e y más realista que la Euclídea porque pone en evidencia la estructura interna del material. Cuando la difusión está gobernada por una ecuación parabólica del tipo $\frac{\partial}{\partial t} = \Delta$, o $\frac{\partial}{\partial t} = \text{div}(A \text{ grad})$ y el operador elíptico del lado derecho tiene un análisis espectral conocido en su dominio, es decir, cuando $\{\lambda_j\}$; $\{\phi_j\}$ son los autovalores y autofunciones del operador elíptico, esa misma distancia puede calcularse así

$$\sqrt{\sum_j e^{2t\lambda_j} |\phi_j(x) - \phi_j(y)|^2}.$$

Esta idea puede trasladarse a cualquier contexto en el que se tenga bien definido un operador elíptico como el Laplaciano. Como ya vimos en el Capítulo 1 este es el caso de los grafos ponderados no dirigidos. En efecto $\Delta = A^{-1}(W - D)$ con la notación de la Sección 1.3. El Corolario 1.3.5 provee dos listas, n autovalores, contados con multiplicidad y n autofunciones: $\lambda_{n-1} \leq \lambda_{n-2} \leq \dots \leq \lambda_1 \leq \lambda_0 = 0$ y $\{\phi_j : j = 0, \dots, n-1\}$, $\phi_0 = \text{cte.}$ que satisfacen

$$\Delta \phi_j = \lambda_j \phi_j ; j = 0, \dots, n-1.$$

Definimos entonces, dado $t > 0$, y dos vértices l y m cualesquiera del grafo,

$$d_t^2(l, m) = \sum_{j=1}^{n-1} e^{2t\lambda_j} |\phi_j(l) - \phi_j(m)|^2.$$

Esta definición provee toda una familia $\{d_t : t > 0\}$ de métricas en el grafo. Consideraremos en otros capítulos propiedades de la métricas que la integran y de continuidad con respecto a t para las distancias en espacios métricos que analizaremos.

Capítulo 3

Dinámica topológica por umbralado de afinidades

3.1. Introducción

En la Sección 2.2 del Capítulo 2, hemos repasado las nociones de conectividad (o conexión) y de componente conexa. El siguiente enunciado contiene un resultado elemental que nos permite definir la componente conexa que contiene a un punto cualquiera de X .

LEMA 3.1.1. *Sea (X, τ) un espacio topológico. Si $\{K_\alpha : \alpha \in A\}$ es una familia arbitraria de conexos tal que $\bigcap_{\alpha \in A} K_\alpha \neq \emptyset$, entonces $\bigcup_{\alpha \in A} K_\alpha$ es conexo.*

DEMOSTRACIÓN. Sea $x \in \bigcap_{\alpha \in A} K_\alpha$. Si $K = \bigcup_{\alpha \in A} K_\alpha$ no fuese conexo, existirían U y V abiertos disjuntos en τ tales que $U \cap K \neq \emptyset$, $V \cap K \neq \emptyset$ y $K \subset U \cup V$. Como $U \cap V = \emptyset$ entonces x pertenece sólo a uno de ellos. Supongamos $x \in U$ y $x \notin V$. Como $V \cap K \neq \emptyset$ existe $\alpha \in A$ tal que $V \cap K_\alpha \neq \emptyset$. Por otra parte $K_\alpha \setminus V \subset U$, puesto que $K_\alpha \subset K \subset U \cup V$. Además $U \cap K_\alpha \supset \{x\}$, entonces $U \cap K_\alpha \neq \emptyset$. Lo que contradice la conectividad de K_α . \square

Este lema nos permite, dado un punto $x \in X$ definir $C(x)$ como la única componente conexa (conexo maximal en X) que contiene a x . En efecto

$$C(x) = \bigcup_{\{C \text{ es un conexo tal que } x \in C\}} C.$$

En este capítulo nos proponemos estudiar la evolución (dinámica) del cardinal del conjunto de las componentes conexas de un grafo no dirigido y ponderado en función de niveles de umbral en la afinidad que determina la relación entre vértices. Usaremos algoritmos en Python para contar las componentes conexas obtenidas y para obtener gráficas del número de componentes conexas como funciones del nivel de umbral. Aplicaremos este análisis al grafo del transporte en AMBA introducido en el Capítulo 1 y a algunas variaciones del mismo que tienen en cuenta otras nociones de vecindad aparte de aquellas

determinadas por el transporte público. Algunos de estos resultados forman del trabajo [AAG23].

3.2. Umbrales de una afinidad y la familia de topologías inducidas

Sea X un conjunto y W una afinidad en X en el sentido de la definición del Capítulo 2. Sea $\lambda > 0$ dado. Definimos otra afinidad en X , W^λ , dada por

$$W^\lambda(x, y) = \begin{cases} +\infty & \text{si } W(x, y) > \lambda \\ W(x, y) & \text{si } W(x, y) \leq \lambda \end{cases}$$

Le llamamos el **umbral en λ de W** . Resulta claro que $W^\lambda(x, y) \geq W(x, y)$ y por lo tanto W^λ también es una afinidad. Por el Teorema 2.4.1 tenemos la topología τ_W inducida por la afinidad W y para cada $\lambda > 0$ la topología $\tau_\lambda := \tau_{W^\lambda}$. Sea $\kappa(\lambda)$ la cantidad de componentes conexas del espacio topológico (X, τ_λ) .

PROPOSICIÓN 3.2.1. Sean X , W , W^λ y $\kappa(\lambda)$ como arriba,

- a) si $0 < \lambda_1 < \lambda_2$, entonces $\tau_{\lambda_1} \subseteq \tau_{\lambda_2} \subseteq \tau_W$;
- b) la función $\kappa(\lambda)$ es no decreciente;
- c) si X es finito, entonces, para λ suficientemente grande se tiene que $\tau_\lambda = \tau_W$ y por consiguiente $\kappa(\lambda)$ es la cantidad de componentes conexas de (X, τ_W) .

DEMOSTRACIÓN. a) Sean $0 < \lambda_1 < \lambda_2$. Sea $U \in \tau_{\lambda_1}$, entonces para todo $x \in U$ existe $\alpha = \alpha(x) > 0$ tal que $E_{\lambda_1}(x, \alpha) \subset U$, donde $E_{\lambda_1}(x, \alpha) = \{y \in X : W^{\lambda_1}(x, y) > \alpha\}$. Pero si y es tal que $W^{\lambda_2}(x, y) > \alpha$, como $W^{\lambda_2}(x, y) \leq W^{\lambda_1}(x, y)$, resulta que $W^{\lambda_1}(x, y) > \alpha$ y entonces $y \in E_{\lambda_1}(x, \alpha)$. Es decir $E_{\lambda_2}(x, \alpha) \subseteq E_{\lambda_1}(x, \alpha)$ y por lo tanto $E_{\lambda_2}(x, \alpha) \subset U$. De aquí que U también está en τ_{λ_2} . Como $W(x, y) \leq W^\lambda(x, y)$ para todo λ , el mismo argumento muestra que $\tau_\lambda \subseteq \tau_W$.

b) Se deduce de a) y de la Proposición 2.2.2.

c) Si X es finito se tiene que también es finito el conjunto de números reales extendidos $W(x, y)$ y por consiguiente $\{W(x, y) : W(x, y) < \infty\}$ tiene máximo. Tomando λ mayor que ese valor máximo se tiene que $W^\lambda \equiv W$. De donde resultan claras las conclusiones en c). □

Notemos que el inciso *c*) no es en general válido en contextos de cardinal arbitrario. En efecto, sea $X = \mathbb{R}$ y $W(x, y) = \frac{1}{|x-y|}$. Entonces, para $\lambda > 0$, tenemos que

$$W^\lambda(x, y) = \begin{cases} \frac{1}{|x-y|} & \text{si } |x-y| \geq \frac{1}{\lambda} \\ \infty & \text{si } |x-y| < \frac{1}{\lambda}. \end{cases}$$

Resulta claro que τ_W es la topología usual ya que los entornos E_λ son intervalos. En cambio, un conjunto abierto no vacío en \mathbb{R} con la topología τ_λ no puede ser otra cosa que el espacio total \mathbb{R} . Veamos esto. Sea $U \subseteq \mathbb{R}$; $U \neq \emptyset$ y $U \in \tau_\lambda$. Sea $x \in U$ entonces existe $\alpha > 0$ tal que $E_\lambda(x, \alpha) \subset U$. Pero

$$E_\lambda(x, \alpha) = \{y \in \mathbb{R} : W^\lambda(x, y) > \alpha\} \supset \left(x - \frac{1}{\lambda}, x + \frac{1}{\lambda}\right).$$

Por lo tanto si $x \in U$, necesariamente un intervalo de la forma $(x - \frac{1}{\lambda}, x + \frac{1}{\lambda})$ está contenido en U . Entonces $U = \mathbb{R}$ y la topología τ_λ es la trivial $\tau_\lambda = \{\emptyset, \mathbb{R}\}$ para todo λ . Pero τ_W es la topología usual.

3.3. Algunos ejemplos sencillos

En situaciones sencillas de grafos ponderados con un número pequeño de vértices se puede ilustrar el comportamiento de la función monótona $\kappa(\lambda)$ en relación con la distribución de los vértices.

En los ejemplos que siguen los conjuntos X que describimos en cada caso serán subconjuntos de \mathbb{R} y la afinidad será siempre la restricción a $X \times X$ de $W(x, y) = \frac{1}{|x-y|}$ donde $|x-y|$ es la distancia usual en \mathbb{R} .

EJEMPLO 3.3.1. Sea $X = \{1, 2, 3, 4, 5, 6\}$ y $W(x, y) = \frac{1}{|x-y|}$. Entonces

$$W = \begin{pmatrix} \infty & 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ 1 & \infty & 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & 1 & \infty & 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{2} & 1 & \infty & 1 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{3} & \frac{1}{2} & 1 & \infty & 1 \\ \frac{1}{5} & \frac{1}{4} & \frac{1}{3} & \frac{1}{2} & 1 & \infty \end{pmatrix}$$

Como para $\lambda \geq 1$ tenemos $W^\lambda = W$, por *c*) de la Proposición 3.2.1, $\kappa(\lambda)$ es igual a la cantidad de componentes conexas de X en τ_W . Recordemos que $\tau_W = \{U \subset X : \text{para todo } i \in U \text{ existe } \alpha > 0 \text{ tal que } E_W(i, \alpha) \subset U\}$, con $E_W(i, \alpha) = \{j : W_{ij} > \alpha\}$. Tomando $\alpha = 2$, tenemos que $E_W(i, 2) = \{i\}$. Por consiguiente cada conjunto unitario en X es abierto. Entonces $\tau_W = \mathcal{P}(X)$ y $\kappa(\lambda) = 6$ si $\lambda \geq 1$. Por otra parte si $\lambda < \frac{1}{5}$, tenemos que $W_{ij}^\lambda = +\infty$, para todo i y todo j . Entonces $\tau^\lambda = \{\emptyset, X\}$. Por consiguiente $\kappa(\lambda) = 1$. Nos resta estudiar el caso de $\frac{1}{5} \leq \lambda < 1$. Si por ejemplo $\frac{1}{2} \leq \lambda < 1$, tenemos que

$$W^\lambda = W^{\frac{1}{2}} = \begin{pmatrix} \infty & \infty & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \infty & \infty & \infty & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \infty & \infty & \infty & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{2} & \infty & \infty & \infty & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{3} & \frac{1}{2} & \infty & \infty & \infty \\ \frac{1}{5} & \frac{1}{4} & \frac{1}{3} & \frac{1}{2} & \infty & \infty \end{pmatrix}$$

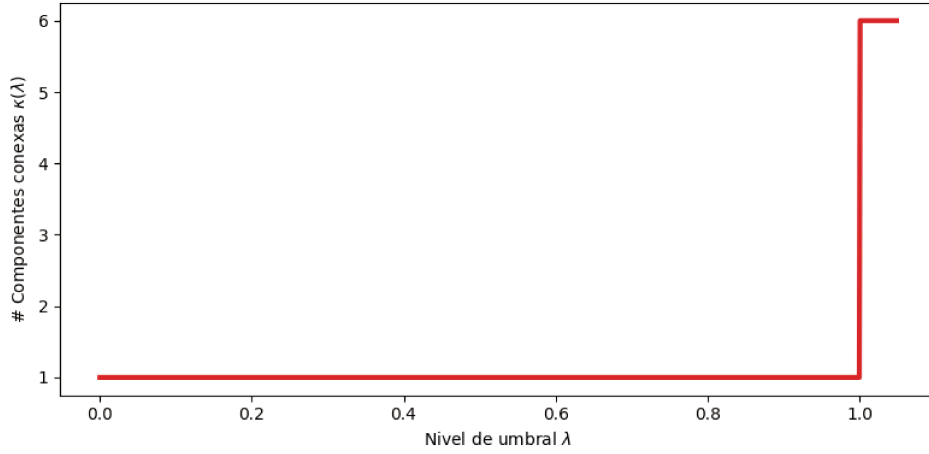
Los entornos de cada punto para $\alpha = 2$ están dados por,

$$E_{\frac{1}{2}}(i, 2) = \begin{cases} \{1, 2\} & \text{si } i = 1; \\ \{i - 1, i, i + 1\} & \text{si } 1 < i < 6; \\ \{6, 5\} & \text{si } i = 6. \end{cases}$$

Entonces $U \in \tau^{\frac{1}{2}}$ si para cada $i \in U$ existe $\alpha \geq 2 : E_{\frac{1}{2}}(i, \alpha) \subset U$. Pero si $1 < i < 6$, entonces $i - 1$ e $i + 1$ también tiene que estar en U . Si $i = 1$, entonces $2 \in U$. Si $i = 6$, entonces $5 \in U$. Entonces $\tau^{\frac{1}{2}} = \{\emptyset, X\}$ y $\kappa(\frac{1}{2}) = 1$. Lo mismo ocurre en los otros sub-intervalos de $[\frac{1}{5}, 1)$ para los distintos valores de λ . Por consiguiente

$$\tau_\lambda = \begin{cases} \{\emptyset, X\} & \text{si } 0 < \lambda < 1 \\ \mathcal{P}(X) & \text{si } \lambda \geq 1. \end{cases}$$

Entonces $\kappa(\lambda)$ se la puede visualizar que tiene la siguiente gráfica en la Figura 3.3.1.

FIGURA 3.3.1. $\kappa(\lambda)$ para el Ejemplo 3.3.1.

Para tener una mejor representación visual de estas gráficas de funciones escalonadas, usaremos las verticales en los puntos de salto, sin perder la perspectiva de que κ es una función que tomá sólo valores enteros.

EJEMPLO 3.3.2. Sea $X = \{1, 2, 3, 11, 12, 13\}$ y, como antes, $W(x, y) = \frac{1}{|x-y|}$. Entonces

$$W = \begin{pmatrix} \infty & 1 & \frac{1}{2} & \frac{1}{10} & \frac{1}{11} & \frac{1}{12} \\ 1 & \infty & 1 & \frac{1}{9} & \frac{1}{10} & \frac{1}{11} \\ \frac{1}{2} & 1 & \infty & \frac{1}{8} & \frac{1}{9} & \frac{1}{10} \\ \frac{1}{10} & \frac{1}{9} & \frac{1}{8} & \infty & 1 & \frac{1}{2} \\ \frac{1}{11} & \frac{1}{10} & \frac{1}{9} & 1 & \infty & 1 \\ \frac{1}{12} & \frac{1}{11} & \frac{1}{10} & \frac{1}{2} & 1 & \infty \end{pmatrix}$$

Para $\lambda \geq 1$ tenemos que $\tau_W = \mathcal{P}(X)$ y $\kappa(\lambda) = 6$. Y para $\lambda < \frac{1}{12}$, la matriz W^λ es la que tiene a $+\infty$ en todas sus entradas. Entonces $\tau_\lambda = \{\emptyset, X\}$ y $\kappa(\lambda) = 1$. Ahora si tomamos $\lambda = \frac{1}{8}$, tenemos que

$$W^\lambda = W^{\frac{1}{8}} = \begin{pmatrix} \infty & \infty & \infty & \frac{1}{10} & \frac{1}{11} & \frac{1}{12} \\ \infty & \infty & \infty & \frac{1}{9} & \frac{1}{10} & \frac{1}{11} \\ \infty & \infty & \infty & \frac{1}{8} & \frac{1}{9} & \frac{1}{10} \\ \frac{1}{10} & \frac{1}{9} & \frac{1}{8} & \infty & \infty & \infty \\ \frac{1}{11} & \frac{1}{10} & \frac{1}{9} & \infty & \infty & \infty \\ \frac{1}{12} & \frac{1}{11} & \frac{1}{10} & \infty & \infty & \infty \end{pmatrix}$$

Entonces para $\alpha \geq 2$ e $i \in X$

$$E_{\frac{1}{8}}(i, 2) = \begin{cases} \{1, 2, 3\} & \text{si } i = 1, 2, 3; \\ \{11, 12, 13\} & \text{si } i = 11, 12, 13 \end{cases}$$

y por lo tanto $\tau_{\frac{1}{8}} = \{\emptyset, X, \{1, 2, 3\}, \{11, 12, 13\}\}$ y $\kappa(\frac{1}{8}) = 2$. Por consiguiente

$$\tau_\lambda = \begin{cases} \{\emptyset, X\} & \text{si } 0 < \lambda < \frac{1}{8} \\ \{\emptyset, X, \{1, 2, 3\}, \{11, 12, 13\}\} & \text{si } \frac{1}{8} \leq \lambda < 1 \\ \mathcal{P}(X) & \text{si } \lambda \geq 1. \end{cases}$$

Vemos que $\kappa(\lambda)$ tiene la siguiente gráfica en la Figura 3.3.2.

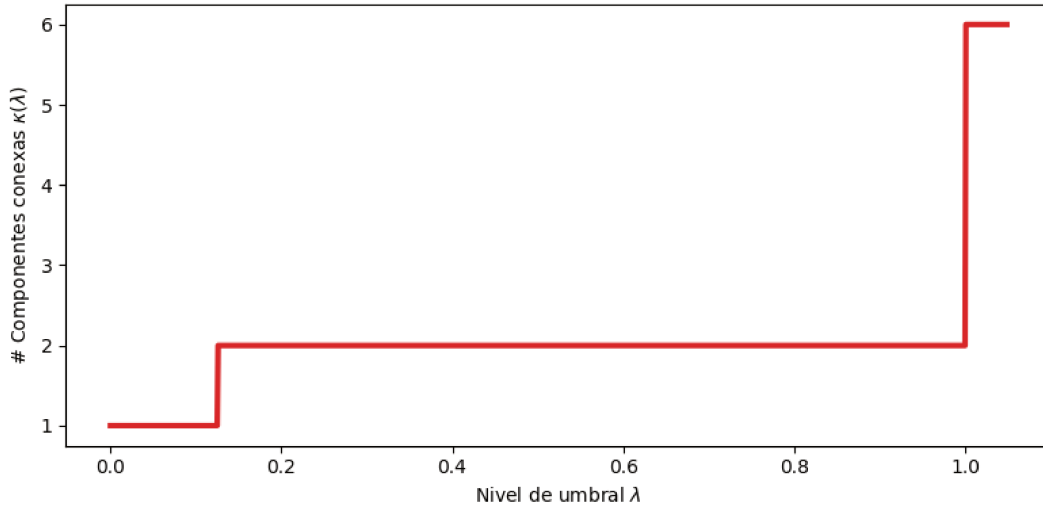


FIGURA 3.3.2. $\kappa(\lambda)$ para el Ejemplo 3.3.2.

EJEMPLO 3.3.3. Sea $X = \{1, 2, 11, 12, 21, 22\}$. Entonces

$$W = \begin{pmatrix} \infty & 1 & \frac{1}{10} & \frac{1}{11} & \frac{1}{20} & \frac{1}{21} \\ 1 & \infty & \frac{1}{9} & \frac{1}{10} & \frac{1}{19} & \frac{1}{20} \\ \frac{1}{10} & \frac{1}{9} & \infty & 1 & \frac{1}{10} & \frac{1}{11} \\ \frac{1}{11} & \frac{1}{10} & 1 & \infty & \frac{1}{9} & \frac{1}{10} \\ \frac{1}{20} & \frac{1}{19} & \frac{1}{10} & \frac{1}{9} & \infty & 1 \\ \frac{1}{21} & \frac{1}{20} & \frac{1}{11} & \frac{1}{10} & 1 & \infty \end{pmatrix}$$

Nuevamente $\tau_W = \mathcal{P}(X)$ y $\kappa(\lambda) = 6$ para $\lambda \geq 1$. Por otra parte si $\lambda < \frac{1}{22}$, de nuevo W^λ es la matriz que tiene a $+\infty$ en todas sus entradas. Entonces $\tau_\lambda = \{\emptyset, X\}$ y $\kappa(\lambda) = 1$.

Tomando $\lambda = \frac{1}{9}$, tenemos que

$$W^\lambda = W^{\frac{1}{9}} = \begin{pmatrix} \infty & \infty & \frac{1}{10} & \frac{1}{11} & \frac{1}{20} & \frac{1}{21} \\ \infty & \infty & \frac{1}{9} & \frac{1}{10} & \frac{1}{19} & \frac{1}{20} \\ \frac{1}{10} & \frac{1}{9} & \infty & \infty & \frac{1}{10} & \frac{1}{11} \\ \frac{1}{11} & \frac{1}{10} & \infty & \infty & \frac{1}{9} & \frac{1}{10} \\ \frac{1}{20} & \frac{1}{19} & \frac{1}{10} & \frac{1}{9} & \infty & \infty \\ \frac{1}{21} & \frac{1}{20} & \frac{1}{11} & \frac{1}{10} & \infty & \infty \end{pmatrix}$$

Entonces para $\alpha \geq 2$ e $i \in X$

$$E_{\frac{1}{9}}(i, 2) = \begin{cases} \{1, 2\} & \text{si } i = 1, 2; \\ \{11, 12\} & \text{si } i = 11, 12; \\ \{21, 22\} & \text{si } i = 21, 22. \end{cases}$$

Y por consiguiente $\tau_{\frac{1}{9}} = \{\emptyset, X, \{1, 2\}, \{11, 12\}, \{21, 22\}, \{1, 2, 11, 12\}, \{1, 2, 21, 22\}, \{11, 12, 21, 22\}\}$ y $\kappa(\frac{1}{9}) = 3$. Por consiguiente

$$\tau_\lambda = \begin{cases} \{\emptyset, X\} & \text{si } 0 < \lambda < \frac{1}{9} \\ \tau_{\frac{1}{9}} & \text{si } \frac{1}{9} \leq \lambda < 1 \\ \mathcal{P}(X) & \text{si } \lambda \geq 1. \end{cases}$$

No es difícil ver que $\kappa(\lambda)$ tiene la gráfica de la Figura 3.3.3.

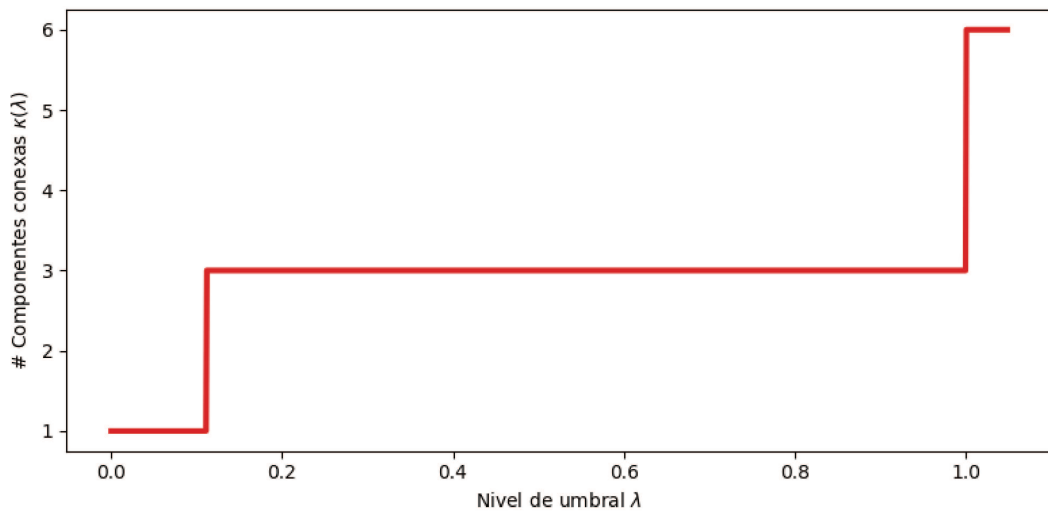
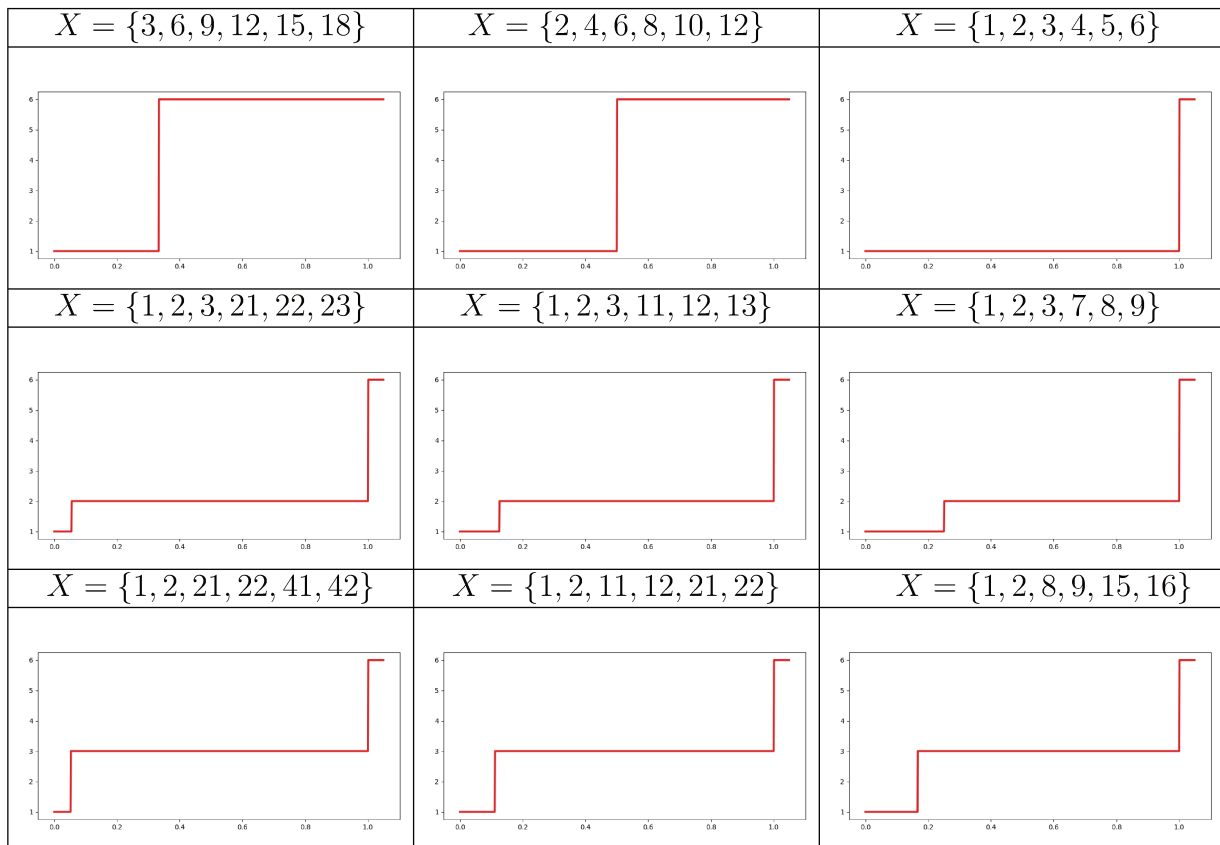


FIGURA 3.3.3. $\kappa(\lambda)$ para el Ejemplo 3.3.3.



CUADRO 1. Ilustración de la función $\kappa(\lambda)$ para distintas separaciones de los clúster.

Los ejemplos anteriores muestran de manera simple que la función $\kappa(\lambda)$ refleja en sus valores la cantidad de agrupamientos que el conjunto muestra en diversas escalas. La tabla que sigue expone cómo esta cantidad de agrupamientos puede manifestarse para distintos valores del umbral cuando los clústeres están separados de manera diferente, la primera fila contiene el Ejemplo 3.3.1 y dos variaciones, se observa que a aumentar la separación de los clústeres el salto se produce para niveles del umbral más bajo, el mismo efecto se muestra para el Ejemplo 3.3.2 segunda fila y el Ejemplo 3.3.3 en la tercera fila de la tabla.

EJEMPLO 3.3.4. Aún en casos sencillos, de los ejemplos anteriores de los seis números naturales puede conseguirse grafos tales que $\kappa(\lambda)$ toma todos los valores entre uno y seis. Un ejemplo concreto se consigue tomando $X = \{2, 5, 6, 11, 22, 24\}$. La grafica de $\kappa(\lambda)$ esta contenida en la siguiente Figura 3.3.4.

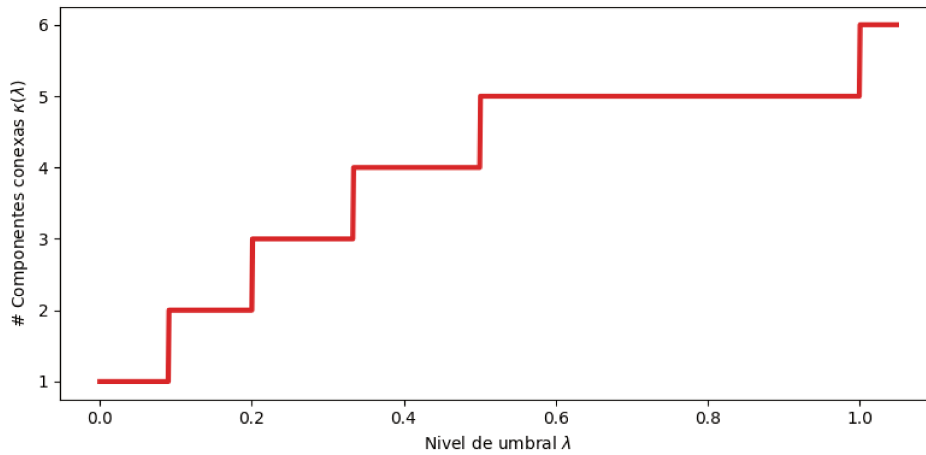


FIGURA 3.3.4. $\kappa(\lambda)$ para el Ejemplo 3.3.4.

3.4. Conectividad topológica y conectividad por caminos de grafos

Sea W una afinidad en $X = \{x_1, \dots, x_n\}$. En este caso W es una matriz simétrica de $n \times n$ tal que $w_{ii} = +\infty$, $i = 1, \dots, n$. Claro que w_{ij} puede valer $+\infty$ en otros pares (i, j) no diagonales, siempre que respete la simetría de la matriz W . A partir de W construimos un grafo simple \mathcal{G}_W de la siguiente manera. Sea $\mathcal{V} = X$ el conjunto de vértices, sea $\mathcal{E}_W = \{\{i, j\} : w_{ij} = +\infty\}$ el conjunto de las aristas. Así, $\mathcal{G}_W = (\mathcal{V}, \mathcal{E}_W)$. La matriz de adyacencia de \mathcal{G}_W está dada por $\mathbf{a}_{ij} = 1$ si $\{i, j\} \in \mathcal{E}_W$ y $\mathbf{a}_{ij} = 0$ si $\{i, j\} \notin \mathcal{E}_W$.

Dado un grafo simple $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ se dice que hay un **camino** que une los vértices con índices i y j , con $i \neq j$, si existen $i = i_1, i_2, \dots, i_k = j$ tales que $k > 1$ y si \mathbf{a} es la matriz de adyacencia de \mathcal{G} , se tiene que $\mathbf{a}_{i_l i_{l+1}} = 1$, $l = 1, \dots, k - 1$. Un subconjunto $\mathcal{V}' \subseteq \mathcal{V}$ es **conexo por caminos** en \mathcal{G} si dados dos vértices cualesquiera en \mathcal{V}' , se pueden conectar por un camino en \mathcal{V}' . Una **componente conexa por caminos** en un grafo \mathcal{G} es un conexo por caminos maximal. En otras palabras si \mathcal{V}^* es componente conexa y $\tilde{\mathcal{V}} \supsetneq \mathcal{V}^*$, entonces $\tilde{\mathcal{V}}$ no es conexo por caminos.

La siguiente proposición será útil para calcular el número de componentes conexas usando el algoritmo de Python `number_connected_components`.

PROPOSICIÓN 3.4.1. Sea W una afinidad dada en el conjunto $X = \{x_1, x_2, \dots, x_n\}$ y sea \mathcal{G}_W el grafo simple inducido por W . Entonces

- i) un subconjunto \tilde{X} de X es conexo con respecto a la topología inducida por la afinidad τ_W en X si y solo si \tilde{X} es conexo por caminos en \mathcal{G}_W ;
- ii) el número de componentes conexas en el sentido topológico en (X, τ_W) coincide con el número de componentes conexas por caminos con respecto al grafo simple \mathcal{G}_W .

DEMOSTRACIÓN. Observamos primero que ii) sigue de i). Para demostrar i), solo necesitamos mostrar que $\mathbf{a}_{ij} = 1$ si y solo si para cualquier conjunto abierto τ_W que contenga a i también contiene a j . Pero $\mathbf{a}_{ij} = 1$ si y solo si $w_{ij} = +\infty$. Tomemos U un conjunto abierto en τ_W tal que $i \in U$. Entonces, existe $\alpha > 0$ tal que $E(i, \alpha) \subset U$. Dado que $E(i, \alpha) = \{k : w_{ik} > \alpha\}$ y $w_{ij} = +\infty$ tenemos que $j \in E(i, \alpha) \subset U$. Por lo tanto $j \in U$. Por otra parte, si j pertenece a cualquier conjunto abierto que contiene a i , entonces $j \in E(i, \alpha)$ para cada $\alpha > 0$. Esto implica que $w_{ij} = +\infty$ y $\mathbf{a}_{ij} = 1$. \square

EJEMPLO 3.4.1. Consideremos el conjunto de los extremos izquierdos de los intervalos que no son excluidos en la cuarta etapa del proceso de la construcción del conjunto de Cantor en el intervalo $[0, 1]$, $X = \{0, \frac{2}{81}, \frac{6}{81}, \frac{8}{81}, \frac{18}{81}, \frac{20}{81}, \frac{24}{81}, \frac{26}{81}, \frac{54}{81}, \frac{56}{81}, \frac{60}{81}, \frac{62}{81}, \frac{72}{81}, \frac{74}{81}, \frac{78}{81}, \frac{80}{81}\}$ la grafica de la función $\kappa(\lambda)$ refleja la clusterización y la separación de los distintos agrupamientos, con $W(x, y) = \frac{1}{|x-y|}$.

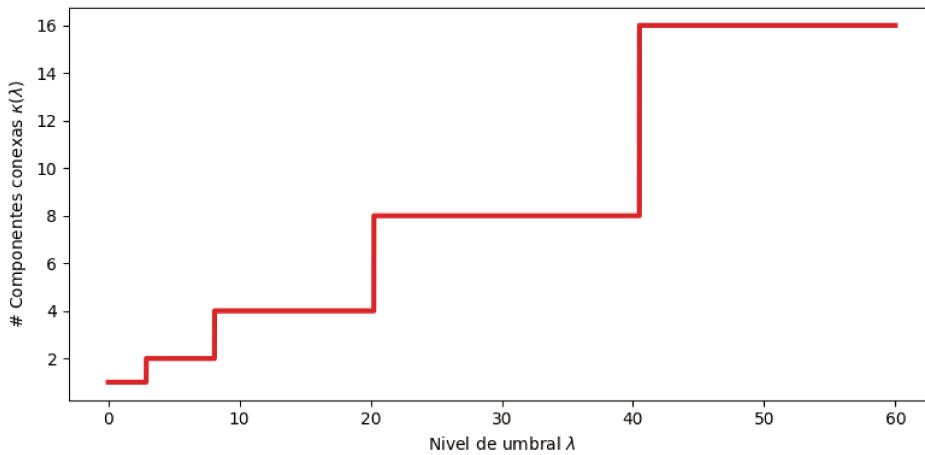


FIGURA 3.4.5. $\kappa(\lambda)$ para una aproximación del conjunto de Cantor.

Se observa que la cantidad de componentes conexas toma sólo los valores 2^i con $i = 0, 1, 2, 3, 4$ y esto se debe a las propiedades de autosimilaridad del conjunto X .

Con el objeto de generar intuiciones relacionadas con las propiedades de concavidad o convexidad de la función $\kappa(\lambda)$ en términos de la distribución geométrica de los vértices cuando usamos la afinidad $W(x, y) = \frac{1}{|x-y|}$, calculamos y representamos tres situaciones especiales. En primer lugar, en la Figura 3.4.6 se representa la función $\kappa(\lambda)$ cuando $X = \{\log_2 i : i = 1, \dots, 20\}$. En la Figura 3.4.7 se representa $\kappa(\lambda)$ para $X = \{\sqrt{i-1} : i = 1, \dots, 20\}$ y en la Figura 3.4.8, $\kappa(\lambda)$ cuando $X = \{20(1 - \frac{1}{i}) : i = 1, \dots, 20\}$. Notamos que las gráficas sólo intentan una descripción cualitativa, ya que los rangos de λ son diferentes. En todos los casos en el uso del algoritmo de Python hemos tomado un paso de 10^{-3} para λ .

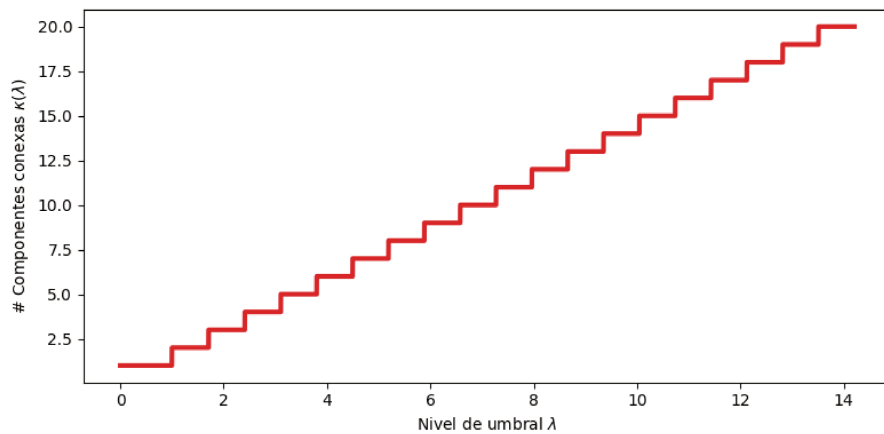


FIGURA 3.4.6. $\{\log_2 i, i = 1, \dots, 20\}$.

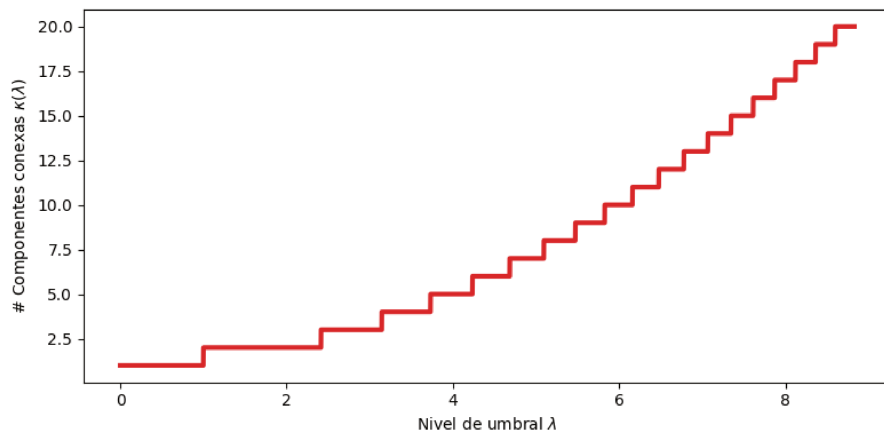


FIGURA 3.4.7. $\{\sqrt{i-1} : i = 1, \dots, 20\}$.

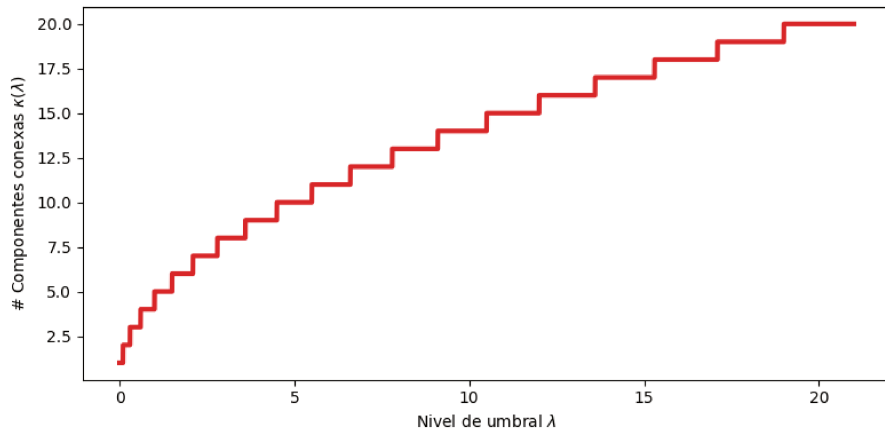


FIGURA 3.4.8. $\{20(1 - \frac{1}{i}) : i = 1, \dots, 20\}$.

Las gráficas de las Figuras 3.4.6, 3.4.7 y 3.4.8 son incomparables por que no hemos normalizado la afinidad en el correspondiente conjunto. Si en cada uno de los tres casos normalizamos la afinidad $\frac{1}{|x-y|}$ dividiendo por el valor máximo que toma en el correspondiente conjunto, entonces obtenemos las curvas de la Figura 3.4.9.

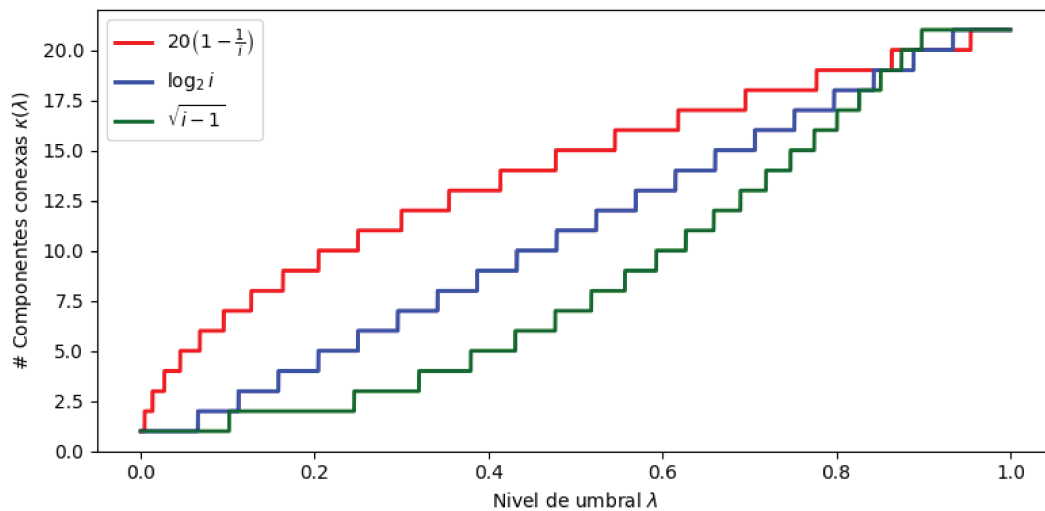


FIGURA 3.4.9. La curva roja corresponde a convergencia de la sucesión de puntos, la verde y la azul a divergencia.

3.5. Dinámica de la conectividad en el transporte público en AMBA

En esta sección vamos a considerar el comportamiento de la función $\kappa(\lambda)$ con el ejemplo del transporte en AMBA presentado en la Sección 1.5 en el Capítulo 1. También consideraremos afinidades de vecindad diferentes de la provista por los datos de SUBE.

Primero vamos a mostrar el comportamiento de $\kappa(\lambda)$ con respecto a la afinidad de la matriz SUBE correspondiente al 4 de marzo de 2020. Sea W_{ij} la matriz de afinidad que se considera en la Sección 1.5 del Capítulo 1. Tomaremos dos tipos de umbrales para la matriz de afinidad que se obtiene de la matriz SUBE tomando el valor $+\infty$ en vez de 0 en la diagonal y algún valor positivo suficientemente chico en los lugares no diagonales ij de la matriz en los que w_{ij} sean ceros.

En la gráfica de la izquierda de la Figura 3.5.10 se muestra la función $\kappa(\lambda)$ tomando como umbrales los máximos no diagonales de cada columna. En la de la derecha se muestra la gráfica de $\kappa(\lambda)$ con umbrales que sean múltiplos enteros de 10^{-3} .

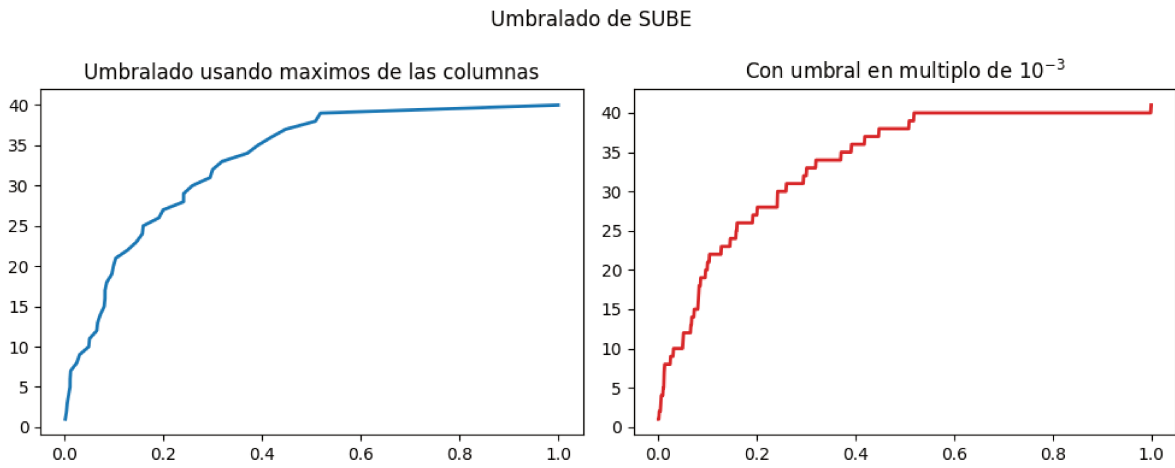


FIGURA 3.5.10. Funciones $\kappa(\lambda)$ para el transporte público.

En los siguientes cuadros, Cuadro 2 y Cuadro 3, representamos la función $\kappa^l(\lambda)$ con $l = 1, \dots, 6$ considerando las seis siguientes matrices entre los 41 distritos de AMBA,

W^1 : $W_{ij}^1 = 1$ si los distritos i y j comparten algún punto de sus fronteras y $W_{ij}^1 = 0$ en caso contrario;

W^2 : $W_{ij}^2 = l_{ij}$ con l_{ij} la longitud de las fronteras compartidas por los distritos i y j ;

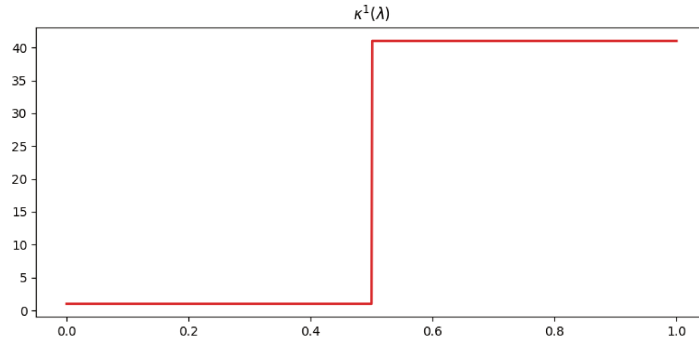
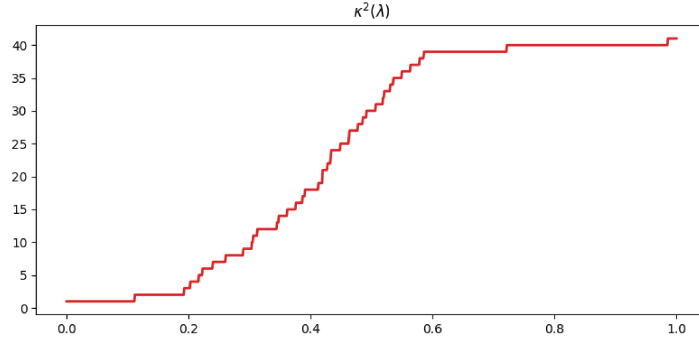
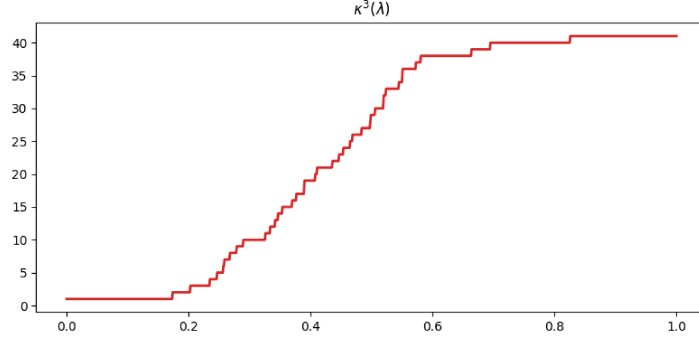
W^3 : $W_{ij}^3 = \frac{1}{d_{ij}}$, donde d_{ij} es la distancia euclídea entre los centros geográficos de los distritos i y j ;

W^4 : es la matriz de SUBE dada en la Sección 1.5, Figura 1.5.7 del Capítulo 1;

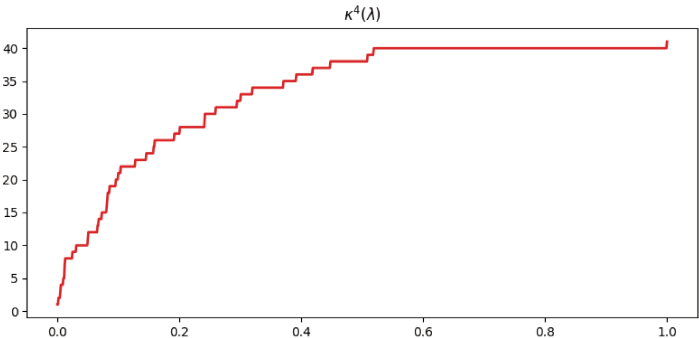
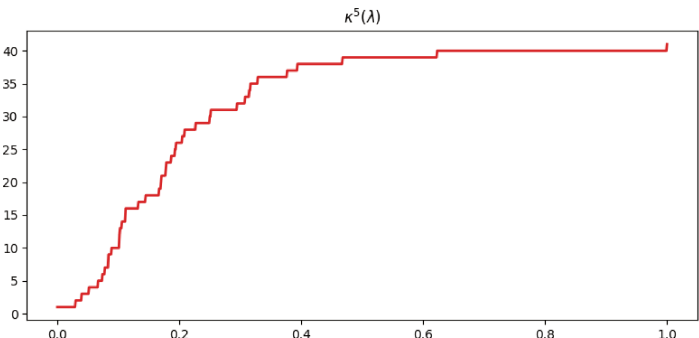
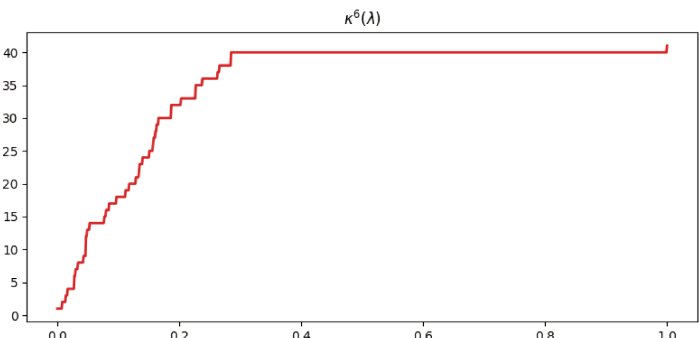
W^5 : $W_{ij}^5 = m_{ij}l_{ij}$, donde m_{ij} es el mínimo de las poblaciones de los distrito i y j y l_{ij} es como W^2 ;

W^6 : $W_{ij}^6 = m_{ij}W_{ij}^1$, donde m_{ij} es como en W^5 .

Todas las matrices han sido normalizadas por su máximo valor (norma infinito) fuera de la diagonal para que los umbrales estén todos en el mismo rango, y las gráficas admitan superposición.

| Afinidad | Función $\kappa^l(\lambda)$ con umbrales múltiplos de 10^{-3} |
|-------------------------------------|--|
| $W^1 : W_{ij}^1$ |  <p>Gráfico de la función $\kappa^1(\lambda)$ que muestra un salto abrupto de 0 a 40 en $\lambda \approx 0.5$.</p> |
| $W^2 : W_{ij}^2 = l_{ij}$ |  <p>Gráfico de la función $\kappa^2(\lambda)$ que muestra una curva escalonada creciente de 0 a 40.</p> |
| $W^3 : W_{ij}^3 = \frac{1}{d_{ij}}$ |  <p>Gráfico de la función $\kappa^3(\lambda)$ que muestra una curva escalonada creciente de 0 a 40.</p> |

CUADRO 2. La función $\kappa(\lambda)$ en AMBA para seis afinidades entre los 41 distritos.

| Afinidad | Función $\kappa^l(\lambda)$ con umbrales múltiplos de 10^{-3} |
|------------------------------------|--|
| $W^4 : W_{ij}^4$ es la matriz SUBE |  |
| $W^5 : W_{ij}^5 = m_{ij}l_{ij}$ |  |
| $W^6 : W_{ij}^6$ |  |

CUADRO 3. La función $\kappa(\lambda)$ en AMBA para seis afinidades entre los 41 distritos.

Una observación elemental esta relacionada con la concavidad (convexidad) “globales”. Mientras que todas las gráficas del Cuadro 3, que involucran variables poblacionales, se muestran globalmente cóncavas, no ocurre lo mismo con las del Cuadro 2 que solo involucran variables geográficas. Si comparamos las graficas de los Cuadros 2 y 3 con la Figura 3.4.9 podríamos postular “convergencia” en los casos que tienen en cuenta variables poblacionales y “divergencia” en los casos que solo hay variables geográficas.

Distancia y Afinidad de Hausdorff

4.1. Introducción

Si bien los resultados del Capítulo 3 muestran estrechas relaciones entre los conceptos de métrica y afinidad, resultados, cuantitativos precisos sólo se tienen bajo condiciones adecuadas de transitividad de la afinidad y algún tipo de triangularidad en la métrica. Por consiguiente el concepto ya clásico de distancia de Hausdorff entre conjuntos de un espacio métrico induce naturalmente un concepto asociado de afinidad entre esos conjuntos, cuando la afinidad es transitiva. En el caso general, esta relación se pierde, pero lo mismo es posible definir una afinidad de tipo Hausdorff entre conjuntos a partir de una afinidad entre los puntos que lo integran. El propósito de este capítulo es doble, por un lado introducir el concepto y las propiedades básicas de afinidad de Hausdorff y por el otro repasar las de la métrica de Hausdorff que serán importantes en el desarrollo posterior. También daremos algunos ejemplos ilustrativos y alguna aplicación de estos conceptos al grafo de SUBE. Recordemos que si (X, d) es un espacio métrico, para A y B subconjuntos de X

$$d(A, B) = \inf\{d(x, y) : x \in A, y \in B\},$$

no define una métrica sobre las partes de X . De hecho, para que $d(A, B) = 0$, basta con que $\bar{A} \cap \bar{B} \neq \emptyset$. Del mismo modo es poco razonable pensar que dos comunidades son muy afines porque comparten un individuo.

En este capítulo exploramos la trasposición de la idea de Hausdorff para definir distancias entre subconjuntos de espacios métricos, para producir una definición natural y útil de afinidad entre subconjuntos de espacios con afinidad. Presentamos también dos algoritmos para calcular, usando Python, distancias y afinidades entre conjuntos.

Sea (X, d) un espacio métrico. Para un subconjunto A de X y un número positivo ϵ se define el conjunto ϵ -paralelo A_ϵ (o ϵ -engordado) por

$$A_\epsilon = \bigcup_{x \in A} B_d(x, \epsilon) = \{x \in X : \exists y \in A \setminus d(x, y) < \epsilon\}.$$

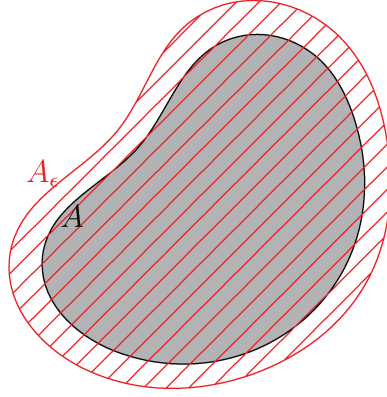


FIGURA 4.1.1. Un conjunto A y el conjunto ϵ -paralelo A_ϵ .

Sea \mathcal{K} la familia de todos los subconjuntos compactos no vacíos en (X, d) ;

$$\mathcal{K} = \{K \subseteq X : K \text{ es compacto y } K \neq \emptyset\}.$$

DEFINICIÓN 4.1.1. Sea (X, d) un espacio métrico. La función d_H definida en $\mathcal{K} \times \mathcal{K}$ por

$$d_H(K^1, K^2) = \inf\{\epsilon > 0 : K^1 \subseteq K_\epsilon^2 \text{ y } K^2 \subseteq K_\epsilon^1\},$$

con K^1 y $K^2 \in \mathcal{K}$ y K_ϵ^i , $i = 1, 2$, sus conjuntos ϵ -paralelos, es llamada **métrica de Hausdorff o distancia de Hausdorff** en \mathcal{K} .

El siguiente resultado es bien conocido y puede encontrarse en [BBI01].

PROPOSICIÓN 4.1.2. *Sea (X, d) un espacio métrico. Sea \mathcal{K} la familia de todos los subconjuntos compactos no vacíos de X . Entonces,*

- (a) d_H esta bien definida en $\mathcal{K} \times \mathcal{K}$;
- (b) d_H es una métrica en \mathcal{K} ;
- (c) si (X, d) es completo entonces (\mathcal{K}, d_H) también es completo.

4.2. Un algoritmo en Python para el cálculo de d_H entre subconjuntos finitos de un espacio métrico

Sea (X, d) un espacio métrico. Si A y B son subconjuntos finitos no vacíos de X , se tiene que A y B están en \mathcal{K} y por consiguiente su distancia de Hausdorff $d_H(A, B)$ está bien definida y es un número real no negativo. La finitud de ambos conjuntos nos permite definir,

- para cada $a \in A$, $\epsilon(a) = \min_{b \in B} d(a, b)$ y $\epsilon = \max_{a \in A} \epsilon(a)$;
- para cada $b \in B$, $\delta(b) = \min_{a \in A} d(a, b)$ y $\delta = \max_{b \in B} \delta(b)$.

PROPOSICIÓN 4.2.1. *Con la notación anterior resulta que $d_H(A, B) = \max\{\epsilon, \delta\}$.*

DEMOSTRACIÓN. Sea $\eta = \max\{\epsilon, \delta\}$.

i) Sea $\eta^+ > \eta$, entonces $\eta^+ > \eta \geq \epsilon \geq \epsilon(a)$, para todo $a \in A$. Entonces para todo $a \in A$ $\min_{b \in B} d(a, b) < \eta^+$ y por consiguiente existe $b \in B$ tal que $d(a, b) < \eta^+$. Es decir para todo $a \in A$ existe $b \in B$ tal que $d(a, b) < \eta^+$. En otras palabras $a \in B_{\eta^+}$ para todo $a \in A$. Entonces

$$(4.1) \quad A \subset B_{\eta^+}.$$

También, siendo $\eta^+ > \eta$, tenemos que $\eta^+ > \eta > \delta \geq \delta(a)$ para todo $b \in B$. Entonces para todo $b \in B$ $\min_{a \in A} d(a, b) < \eta^+$. Es decir, para todo $b \in B$ existe $a \in A$ tal que $d(a, b) < \eta^+$. En otras palabras $b \in A_{\eta^+}$ para todo $b \in B$. Entonces

$$(4.2) \quad B \subset A_{\eta^+}.$$

De (4.1) y (4.2) se deduce que $d_H(A, B) \leq \eta^+$ para todo $\eta^+ > \eta$. En particular $d_H(A, B) \leq \eta$.

ii) Para probar que $d_H(A, B) \leq \eta$, basta entonces verificar que si $0 < \eta^- < \eta$ se tiene que al menos una de las dos inclusiones $A \subset B_{\eta^-}$ o $B \subset A_{\eta^-}$ es imposible. Supongamos entonces que $0 < \eta^- < \eta = \max\{\epsilon, \delta\}$ y que $A \subset B_{\eta^-}$. Entonces para todo $a \in A$ existe $b \in B$ tal que $d(a, b) < \eta^- < \eta$. De aquí que para todo $a \in A$, $\epsilon(a) = \min_{b \in B} d(a, b) < \eta^- < \eta$ y entonces $\epsilon = \max_{a \in A} \epsilon(a) < \eta^- < \eta$. Esto significa que $\eta = \max\{\epsilon, \delta\} = \delta$. De aquí que $\delta = \eta > \eta^-$. Es decir que $\max_{b \in B} \delta(b) > \eta^-$, por consiguiente existe $b \in B$ tal que $\delta(b) > \eta^-$.

Pero entonces $\min_{a \in A} d(a, b) = \delta(b) > \eta^-$. De aquí que existe $b \in B$ tal que para todo $a \in A$ se tiene que $d(a, b) > \eta^-$. Por consiguiente $b \notin A_{\eta^-}$ y $B \not\subset A_{\eta^-}$. \square

Este resultado conduce naturalmente al siguiente algoritmo;

1. $a \in A$ calcular $\epsilon(a) = \min_{b \in B} d(a, b)$;
2. calcular $\epsilon = \max_{a \in A} \epsilon(a)$;
3. $b \in B$ calcular $\delta(b) = \min_{a \in A} d(a, b)$;
4. calcular $\delta = \max_{b \in B} \delta(b)$;
5. calcular $\max\{\epsilon, \delta\} = d_H(A, B)$.

Este algoritmo en términos de scripts de Python luce así;

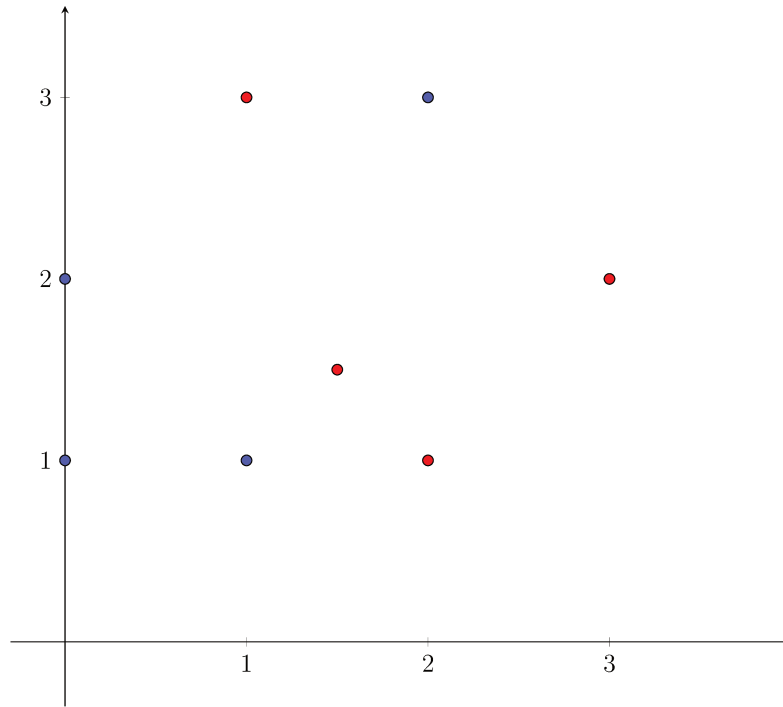
```
def retorna_distancia(A, B, d):
    epsilon_a = []
    for a in A:
        dis_a_b = []
        for b in B:
            dis_a_b.append(d[a][b])
        epsilon_a.append(min(dis_a_b))
    epsilon = max(epsilon_a)

    delta_b = []
    for b in B:
        dis_b_a = []
        for a in A:
            dis_b_a.append(d[b][a])
        delta_b.append(min(dis_b_a))
    delta = max(delta_b)

    d_H(A,B) = max(epsilon, delta)
    return d_H(A,B)
```

En particular nos permite calcular y estimar distancias entre subconjuntos finitos de algún espacio métrico o entre subconjuntos de vértices de grafos métricos. Ilustramos esto en los ejemplos siguientes.

EJEMPLO 4.2.1. Sean $A = \{(0, 1), (0, 2), (2, 3), (1, 1)\}$ y $B = \{(2, 1), (3, 2), (\frac{3}{2}, \frac{3}{2}), (1, 3)\}$ en \mathbb{R}^2 . En la Figura 4.2.2 A se representa con color azul en la y B con color rojo con la métrica usual de \mathbb{R}^2 .

FIGURA 4.2.2. $d_H(A; B) \approx 1.5811388$.

La distancia de Hausdorff entre los conjuntos A y B es $d_H(A, B) \approx 1.5811388$.

EJEMPLO 4.2.2. Consideremos una partición guiada por la vecindad geográfica del conjunto de los 41 distritos del AMBA en 16 regiones que son subconjuntos del conjunto de los números naturales en el intervalo $[1, 41]$. Nos guiaremos por el mapa del Capítulo 1. Así, proponemos la partición de AMBA (A) dada por $A = \bigcup_{j=1}^{16} A_j$; donde

$$A_1 = \{28, 3, 12\} = \{\text{Zárate, Campana, San Fernando}\};$$

$$A_2 = \{11, 4, 27\} = \{\text{Pilar, Escobar, Tigre}\};$$

$$A_3 = \{32, 9\} = \{\text{Exaltación de la Cruz, Luján}\};$$

$$A_4 = \{33, 6, 23\} = \{\text{General Las Heras, General Rodríguez, Marcos Paz}\};$$

$$A_5 = \{17, 13, 16\} = \{\text{Cañuelas, San Vicente, Brandsen}\};$$

$$A_6 = \{22, 31, 2\} = \{\text{La Plata, Ensenada, Berisso}\};$$

$$A_7 = \{29, 19, 25, 1\} = \{\text{Berazategui, Florencio Varela, Presidente Perón, Almirante Brown}\};$$

$$A_8 = \{8, 15, 39\} = \{\text{Lanús, Avellaneda, Quilmes}\};$$

$$A_9 = \{36, 18, 5\} = \{\text{Lomas de Zamora, Esteban Echeverría, Ezeiza}\};$$

$$A_{10} = \{35\} = \{\text{La Matanza}\};$$

$$A_{11} = \{30\} = \{\text{CABA}\};$$

$$A_{12} = \{20, 41\} = \{\text{General San Martín, Tres de Febrero}\};$$

$$A_{13} = \{7, 34, 38\} = \{\text{Ituzaingó, Hurlingham, Morón}\};$$

$$A_{14} = \{24, 10\} = \{\text{Merlo, Moreno}\};$$

$$A_{15} = \{21, 37, 40\} = \{\text{José C. Paz, Malvinas Argentinas, San Miguel}\};$$

$$A_{16} = \{26, 14\} = \{\text{San Isidro, Vicente López}\}.$$

Recordemos que en el Capítulo 2 en la Sección 2.7 introdujimos la distancia difusiva dada por

$$d_t^2(l, m) = \sum_{j=1}^{n-1} e^{2t\lambda_j} |\phi_j(l) - \phi_j(m)|^2,$$

donde $\{\lambda_j\}$, $\{\phi_j\}$ es la resolución espectral del Laplaciano para la matriz SUBE M dada en la Figura 1.5.7. Entonces la métrica difusiva para $t = 100$ en el conjunto $A = \{1, \dots, 41\}$ es la siguiente

$$d_{100}^2(i, j) = \sum_{l=1}^{41} e^{-2 \cdot 100 \lambda_l} |\varphi_l(i) - \varphi_l(j)|^2.$$

Con estos insumos resulta la matriz $d_{100,H}(m, n)$ con $1 \leq m \leq 16$, $1 \leq n \leq 16$ de distancias de Hausdorff mutuas entre las 16 partes A_m en las que se particionó AMBA. Esta matriz está dada por

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|----|---------|---------|---------|---------|---------|---------|---------|----------|---------|----------|----------|----------|---------|----------|---------|---------|
| 0 | 0 | 0.94844 | 1.2649 | 1.1143 | 1.1645 | 0.99448 | 0.94378 | 0.94358 | 0.94276 | 0.94085 | 0.9398 | 0.94365 | 0.9447 | 0.94318 | 0.94929 | 0.94182 |
| 1 | 0.94844 | 0 | 0.97854 | 1.0084 | 0.9954 | 0.85201 | 0.55407 | 0.36346 | 0.43261 | 0.34899 | 0.34249 | 0.34998 | 0.40751 | 0.33994 | 0.32166 | 0.32298 |
| 2 | 1.2649 | 0.97854 | 0 | 1.3089 | 1.299 | 1.1946 | 1.015 | 0.97405 | 0.97328 | 0.97134 | 0.97052 | 0.9739 | 0.97478 | 0.97223 | 0.97836 | 0.97187 |
| 3 | 1.1143 | 1.0084 | 1.3089 | 0 | 1.2119 | 1.0498 | 1.0026 | 1.0025 | 1.0017 | 0.99969 | 0.99934 | 1.0024 | 1.0027 | 1.0012 | 1.0078 | 1.0012 |
| 4 | 1.1645 | 0.9954 | 1.299 | 1.2119 | 0 | 1.0861 | 0.98789 | 0.98853 | 0.98754 | 0.98663 | 0.98586 | 0.98935 | 0.9902 | 0.98897 | 0.99493 | 0.98796 |
| 5 | 0.99448 | 0.85201 | 1.1946 | 1.0498 | 1.0861 | 0 | 0.84123 | 0.83901 | 0.84043 | 0.83934 | 0.83699 | 0.84351 | 0.84484 | 0.8439 | 0.8514 | 0.84174 |
| 6 | 0.94378 | 0.55407 | 1.015 | 1.0026 | 0.98789 | 0.84123 | 0 | 0.51273 | 0.50723 | 0.52292 | 0.51685 | 0.53244 | 0.53699 | 0.53775 | 0.55242 | 0.53062 |
| 7 | 0.94358 | 0.36346 | 0.97405 | 1.0025 | 0.98853 | 0.83901 | 0.51273 | 0 | 0.37883 | 0.1006 | 0.087027 | 0.12831 | 0.38461 | 0.14819 | 0.20753 | 0.14787 |
| 8 | 0.94276 | 0.43261 | 0.97328 | 1.0017 | 0.98754 | 0.84043 | 0.50723 | 0.37883 | 0 | 0.38576 | 0.3841 | 0.40212 | 0.40724 | 0.40895 | 0.42999 | 0.40144 |
| 9 | 0.94085 | 0.34899 | 0.97134 | 0.99969 | 0.98663 | 0.83934 | 0.52292 | 0.1006 | 0.38576 | 0 | 0.025545 | 0.086013 | 0.35583 | 0.078092 | 0.18238 | 0.11964 |
| 10 | 0.9398 | 0.34249 | 0.97052 | 0.99934 | 0.98586 | 0.83699 | 0.51685 | 0.087027 | 0.3841 | 0.025545 | 0 | 0.087192 | 0.36667 | 0.091312 | 0.17521 | 0.11228 |
| 11 | 0.94365 | 0.34998 | 0.9739 | 1.0024 | 0.98935 | 0.84351 | 0.53244 | 0.12831 | 0.40212 | 0.086013 | 0.087192 | 0 | 0.36893 | 0.11028 | 0.18185 | 0.11013 |
| 12 | 0.9447 | 0.40751 | 0.97478 | 1.0027 | 0.9902 | 0.84484 | 0.53699 | 0.38461 | 0.40724 | 0.35583 | 0.36667 | 0.36893 | 0 | 0.35595 | 0.3918 | 0.37886 |
| 13 | 0.94318 | 0.33994 | 0.97223 | 1.0012 | 0.98897 | 0.8439 | 0.53775 | 0.14819 | 0.40895 | 0.078092 | 0.091312 | 0.11028 | 0.35595 | 0 | 0.16716 | 0.12557 |
| 14 | 0.94929 | 0.32166 | 0.97836 | 1.0078 | 0.99493 | 0.8514 | 0.55242 | 0.20753 | 0.42999 | 0.18238 | 0.17521 | 0.18185 | 0.3918 | 0.16716 | 0 | 0.15158 |
| 15 | 0.94182 | 0.32298 | 0.97187 | 1.0012 | 0.98796 | 0.84174 | 0.53062 | 0.14787 | 0.40144 | 0.11964 | 0.11228 | 0.11013 | 0.37886 | 0.12557 | 0.15158 | 0 |

FIGURA 4.2.3. Distancias difusivas de Hausdorff entre conglomerados de AMBA.

Notemos por ejemplo que $A_{10} = \{\text{La Matanza}\}$ y $A_{11} = \{\text{CABA}\}$ son, naturalmente, los más cercanos con respecto a $d_{100,H}$. En la Figura 4.2.3, esto corresponde a la intersección de la fila 9 y la columna 10, puesto que en el formato de Python enumera de 0 a 15.

4.3. Afinidad de Hausdorff entre subconjuntos de un espacio con afinidad

Sea (X, W) un espacio con afinidad. Dados $x \in X$ y $\alpha > 0$ consideramos, como antes, el conjunto $E(x, \alpha) = \{y \in X : W(x, y) > \alpha\}$. Recordemos que una afinidad W en X es una función positiva en $X \times X$ que vale $+\infty$ sobre la diagonal. Supondremos de aquí en más que W es **simétrica**. Es decir $W(x, y) = W(y, x)$. Los resultados del Capítulo 2 nos muestran cómo una afinidad en X genera una topología en X a través del sistema de entornos basados en $\mathcal{E}_x = \{E(x, \alpha) : \alpha > 0\}$.

Si $A \subset X$ es distinto del vacío y $\alpha > 0$ definimos el α -entorno de afinidad de A por

$$\begin{aligned} E(A, \alpha) &= \{y \in X : W(y, x) > \alpha \text{ para algún } x \in A\} \\ &= \bigcup_{x \in A} E(x, \alpha). \end{aligned}$$

LEMA 4.3.1. *Si $x_0 \in X$ y $A \subset X$ entonces $A \subset \bigcup_{n \geq 1} E(x_0, \frac{1}{n})$.*

DEMOSTRACIÓN. Sea $y \in A$, entonces $W(x_0, y) > 0$. Entonces existe $n \geq 1$ tal que $W(x_0, y) > \frac{1}{n}$ y por lo tanto $y \in E(x_0, \frac{1}{n})$. \square

Dados dos subconjuntos A y B de X definimos

$$W_H(A, B) = \sup\{\alpha > 0 : A \subseteq E(B, \alpha) \text{ y } B \subseteq E(A, \alpha)\}.$$

En principio la función W_H así definida podría no ser una afinidad en $\mathcal{P}(X)$ puesto que podría ser nula para algún par de conjuntos no vacíos A y B . No obstante la función W_H satisface las siguientes propiedades

PROPOSICIÓN 4.3.2.

- a) $W_H(A, A) = +\infty$, para todo $A \subset X$;
- b) $W_H(A, B) \geq 0$, para todos A y $B \in \mathcal{P}(X)$;
- c) $W_H(A, B) = W_H(B, A)$, para todo $A \subset X$ y todo $B \subset X$;

d) Si $\mathcal{P}_f(X)$ denotamos la familia de las partes finitas de X entonces W_H es una afinidad en $\mathcal{P}_f(X)$.

DEMOSTRACIÓN. b) y c) son claras de la definición. a) se deduce de $A \subset E(A, \alpha)$ para todo $\alpha > 0$. Probemos d). Basta ver que $W_H(A, B) > 0$ para toda elección de A y B subconjuntos finitos distintos de X . Sea $a \in A$ por el Lema 4.3.1 se tiene que $B \subset \bigcup_{n \geq 1} E(a, \frac{1}{n})$, pero como B es finito tendremos que $B \subset \bigcup_{1 \leq n \leq m(a)} E(a, \frac{1}{n})$ donde $m(a)$ dependerá de $a \in A$. Sea $M = \max_{a \in A} m(a)$, entonces $B \subset \bigcup_{1 \leq n \leq M} E(a, \frac{1}{n})$ para todo $a \in A$. Por consiguiente $B \subset E(A, \frac{1}{M})$. Un razonamiento análogo muestra que $A \subset E(B, \frac{1}{K})$ para algún K suficientemente grande. De aquí que $W_H(A, B) \geq \min\{\frac{1}{M}, \frac{1}{K}\} > 0$. \square

4.4. Un algoritmo en Python para el cálculo de W_H entre subconjuntos finitos de un espacio con afinidad

Sea (X, W) un espacio con afinidad. Considerando $\mathcal{P}_f(X)$ la familia de las partes finitas de X entonces por la Proposición 4.3.2 W_H es una afinidad en $\mathcal{P}_f(X)$. La finitud de ambos conjuntos A y B nos permiten un cálculo explícito de la afinidad de Hausdorff $W_H(A, B)$,

- para cada $a \in A$, definimos $\epsilon^*(a) = \max_{b \in B} W(a, b)$ y luego $\epsilon^* = \min_{a \in A} \epsilon^*(a)$;
- para cada $b \in B$, definimos $\delta^*(b) = \max_{a \in A} W(a, b)$ y luego $\delta^* = \min_{b \in B} \delta^*(b)$.

PROPOSICIÓN 4.4.1. Con la notación anterior resulta que $W_H(A, B) = \min\{\epsilon^*, \delta^*\}$

La demostración sigue el esquema de la demostración de la Proposición 4.2.1.

Este resultado conduce naturalmente al siguiente algoritmo;

1. $a \in A$ calcular $\epsilon^*(a) = \max_{b \in B} W(a, b)$;
2. calcular $\epsilon^* = \min_{a \in A} \epsilon^*(a)$;
3. $b \in B$ calcular $\delta^*(b) = \max_{a \in A} W(a, b)$;
4. calcular $\delta^* = \min_{b \in B} \delta^*(b)$;
5. calcular $\min\{\epsilon^*, \delta^*\} = W_H(A, B)$.

Este algoritmo en términos de scripts de Python luce así;

```
def retorna_distancia(A, B, W):
    epsilon_hat_a = []
```

```

for a in A:
    afin_a_b = []
    for b in B:
        afin_a_b.append(W[a][b])
    epsilon^*_a.append(max(afin_a_b))
epsilon^* = min(epsilon^*_a)

```

```

delta^* = []
for b in B:
    afin_b_a = []
    for a in A:
        afin_b_a.append(W[b][a])
    delta^*_b.append(max(afin_b_a))
delta^* = min(delta^*_b)

```

```

WH(A,B) = min(epsilon^*, delta^*)
return WH(A,B)

```

Volvemos a la partición $A = \bigcup_{j=1}^{16} A_j$ del AMBA de la Sección 4.2 y aplicamos este algoritmo al cálculo de la correspondiente W_H . Adecuamos W_H para que sea una matriz de afinidad tomando el doble del valor máximo en la diagonal en vez de cero y la mitad del valor mínimo en los lugares no diagonales ij de la matriz en los que w_{ij} sean ceros.

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0 | 176442 | 309 | 31 | 0.5 | 0.5 | 1 | 5 | 35 | 16 | 197 | 5958 | 98 | 28 | 74 | 424 | 2223 |
| 1 | 309 | 176442 | 266 | 0.5 | 0.5 | 1 | 20 | 66 | 22 | 519 | 12837 | 441 | 65 | 277 | 3068 | 4845 |
| 2 | 31 | 266 | 176442 | 1 | 0.5 | 0.5 | 0.5 | 3 | 1 | 38 | 949 | 10 | 20 | 126 | 15 | 16 |
| 3 | 0.5 | 0.5 | 1 | 176442 | 0.5 | 0.5 | 1 | 4 | 4 | 646 | 2198 | 98 | 130 | 2196 | 66 | 115 |
| 4 | 0.5 | 0.5 | 0.5 | 0.5 | 176442 | 2 | 39 | 104 | 321 | 954 | 2679 | 8 | 3 | 9 | 0.5 | 9 |
| 5 | 1 | 1 | 0.5 | 0.5 | 2 | 176442 | 58 | 264 | 44 | 199 | 8465 | 43 | 11 | 40 | 17 | 55 |
| 6 | 5 | 20 | 0.5 | 1 | 39 | 58 | 176442 | 3745 | 808 | 1048 | 34539 | 135 | 16 | 54 | 22 | 184 |
| 7 | 35 | 66 | 3 | 4 | 104 | 264 | 3745 | 176442 | 1021 | 971 | 39521 | 146 | 31 | 115 | 45 | 165 |
| 8 | 16 | 22 | 1 | 4 | 321 | 44 | 808 | 1021 | 176442 | 3783 | 45769 | 553 | 78 | 214 | 86 | 506 |
| 9 | 197 | 519 | 38 | 646 | 954 | 199 | 1048 | 971 | 3783 | 176442 | 88221 | 1512 | 1090 | 2727 | 176 | 1017 |
| 10 | 5958 | 12837 | 949 | 2198 | 2679 | 8465 | 34539 | 39521 | 45769 | 88221 | 176442 | 26014 | 6070 | 22871 | 8749 | 17645 |
| 11 | 98 | 441 | 10 | 98 | 8 | 43 | 135 | 146 | 553 | 1512 | 26014 | 176442 | 484 | 1672 | 822 | 6059 |
| 12 | 28 | 65 | 20 | 130 | 3 | 11 | 16 | 31 | 78 | 1090 | 6070 | 484 | 176442 | 5152 | 291 | 265 |
| 13 | 74 | 277 | 126 | 2196 | 9 | 40 | 54 | 115 | 214 | 2727 | 22871 | 1672 | 5152 | 176442 | 1309 | 1666 |
| 14 | 424 | 3068 | 15 | 66 | 0.5 | 17 | 22 | 45 | 86 | 176 | 8749 | 822 | 291 | 1309 | 176442 | 3462 |
| 15 | 2223 | 4845 | 16 | 115 | 9 | 55 | 184 | 165 | 506 | 1017 | 17645 | 6059 | 265 | 1666 | 3462 | 176442 |

FIGURA 4.4.4. La matriz de afinidad de Hausdorff entre los 16 conglomerados del AMBA.

En la Figura 4.2.3 vimos que el distrito La Matanza con el distrito CABA son los mas cercanos con $d_{100,H}$. En consonancia con esa observación, este nuevo análisis con el mismo

ejemplo de los distritos A_{10} y A_{11} podemos notar que en la Figura 4.4.4 se visualiza que La Matanza con CABA son los más afines.

Cabe preguntarse cuán distintas son las matrices $d_{100,H}(m,n)$ de la Sección 4.2 con alguna $d_{100}^{W_H}(m,n)$ que se obtiene de la matriz W_H vía el análisis espectral de su Laplaciano. Otra vez podemos hacer el cálculo efectivo en Python. Consideramos entonces el Laplaciano provisto por la matriz en la Figura 4.4.4 con los algoritmos de autovalores y autofunciones del Capítulo 1. Ilustramos en la Figura 4.4.5 los autovalores y sus correspondientes autofunciones.

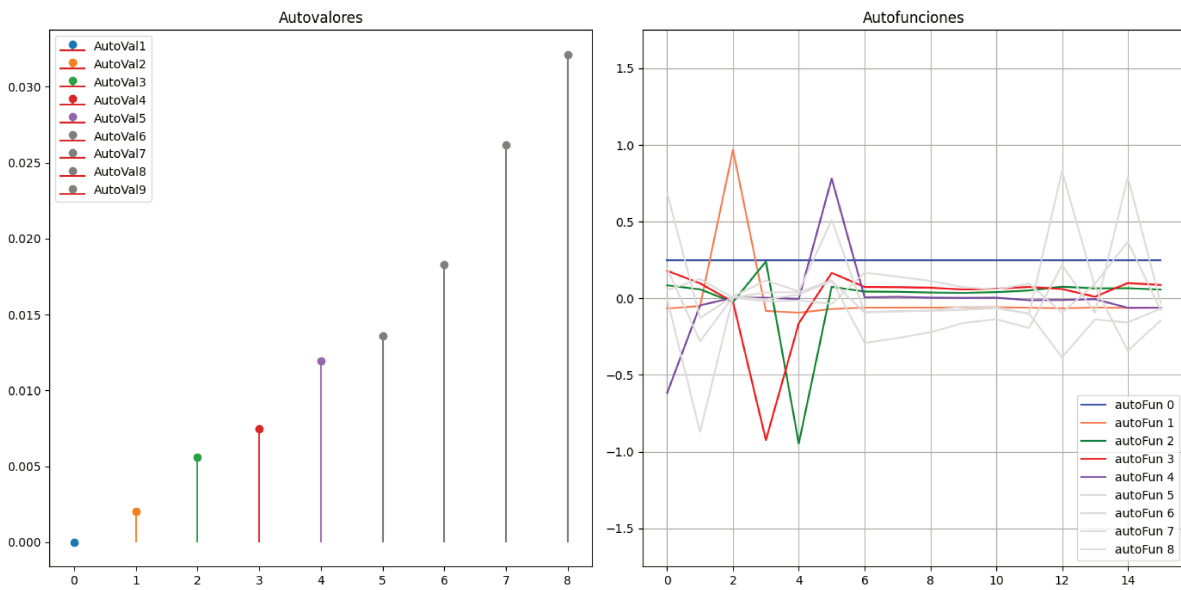


FIGURA 4.4.5. Primeros autovalores y autofunciones de afinidad de Hausdorff entre los 16 conglomerados de AMBA.

Con estos autovalores y autofunciones podemos calcular la distancia difusiva asociada a la matriz W_H . Para $t = 100$, $d_{100}^{W_H}(m,n)$ está dada por

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|----|---------|----------|---------|---------|---------|---------|-----------|-----------|-----------|-----------|-----------|----------|---------|----------|----------|----------|
| 0 | 0 | 0.28675 | 0.88827 | 0.5829 | 0.65966 | 0.42506 | 0.28453 | 0.28369 | 0.28273 | 0.28148 | 0.27911 | 0.28193 | 0.35027 | 0.29315 | 0.29491 | 0.26579 |
| 1 | 0.28675 | 0 | 0.83633 | 0.50345 | 0.59265 | 0.30943 | 0.064727 | 0.064169 | 0.063265 | 0.064483 | 0.062704 | 0.057581 | 0.19393 | 0.08336 | 0.071582 | 0.045183 |
| 2 | 0.88827 | 0.83633 | 0 | 0.96888 | 1.0148 | 0.89339 | 0.84103 | 0.8406 | 0.84012 | 0.83889 | 0.83793 | 0.8415 | 0.86094 | 0.84009 | 0.84831 | 0.84187 |
| 3 | 0.5829 | 0.50345 | 0.96888 | 0 | 0.76787 | 0.58486 | 0.49058 | 0.48988 | 0.48865 | 0.48336 | 0.48473 | 0.48985 | 0.51171 | 0.45984 | 0.50735 | 0.49431 |
| 4 | 0.65966 | 0.59265 | 1.0148 | 0.76787 | 0 | 0.65982 | 0.57889 | 0.57764 | 0.57435 | 0.57182 | 0.57483 | 0.58277 | 0.61736 | 0.58585 | 0.5998 | 0.58711 |
| 5 | 0.42506 | 0.30943 | 0.89339 | 0.58486 | 0.65982 | 0 | 0.2877 | 0.28578 | 0.28714 | 0.28676 | 0.28444 | 0.29313 | 0.35478 | 0.30025 | 0.32138 | 0.30003 |
| 6 | 0.28453 | 0.064727 | 0.84103 | 0.49058 | 0.57889 | 0.2877 | 0 | 0.0049629 | 0.0089406 | 0.016262 | 0.016757 | 0.015592 | 0.17177 | 0.047316 | 0.094587 | 0.034375 |
| 7 | 0.28369 | 0.064169 | 0.8406 | 0.48988 | 0.57764 | 0.28578 | 0.0049629 | 0 | 0.0055987 | 0.012998 | 0.013207 | 0.014686 | 0.17115 | 0.046201 | 0.093917 | 0.03337 |
| 8 | 0.28273 | 0.063265 | 0.84012 | 0.48865 | 0.57435 | 0.28714 | 0.0089406 | 0.0055987 | 0 | 0.0085273 | 0.0096071 | 0.013908 | 0.16982 | 0.044333 | 0.092681 | 0.031993 |
| 9 | 0.28148 | 0.064483 | 0.83889 | 0.48336 | 0.57182 | 0.28676 | 0.016262 | 0.012998 | 0.0085273 | 0 | 0.005049 | 0.017809 | 0.16671 | 0.039361 | 0.092832 | 0.033226 |
| 10 | 0.27911 | 0.062704 | 0.83793 | 0.48473 | 0.57483 | 0.28444 | 0.016757 | 0.013207 | 0.0096071 | 0.005049 | 0 | 0.017252 | 0.16773 | 0.040973 | 0.091062 | 0.030673 |
| 11 | 0.28193 | 0.057581 | 0.8415 | 0.48985 | 0.58277 | 0.29313 | 0.015592 | 0.014686 | 0.013908 | 0.017809 | 0.017252 | 0 | 0.16953 | 0.044534 | 0.086575 | 0.024432 |
| 12 | 0.35027 | 0.19393 | 0.86094 | 0.51171 | 0.61736 | 0.35478 | 0.17177 | 0.17115 | 0.16982 | 0.16671 | 0.16773 | 0.16953 | 0 | 0.14083 | 0.20324 | 0.17874 |
| 13 | 0.29315 | 0.08336 | 0.84009 | 0.45984 | 0.58585 | 0.30025 | 0.047316 | 0.046201 | 0.044333 | 0.039361 | 0.040973 | 0.044534 | 0.14083 | 0 | 0.10425 | 0.057685 |
| 14 | 0.29491 | 0.071582 | 0.84831 | 0.50735 | 0.5998 | 0.32138 | 0.094587 | 0.093917 | 0.092681 | 0.092832 | 0.091062 | 0.086575 | 0.20324 | 0.10425 | 0 | 0.074635 |
| 15 | 0.26579 | 0.045183 | 0.84187 | 0.49431 | 0.58711 | 0.30003 | 0.034375 | 0.03337 | 0.031993 | 0.033226 | 0.030673 | 0.024432 | 0.17874 | 0.057685 | 0.074635 | 0 |

FIGURA 4.4.6. Distancia difusiva entre los 16 conglomerados para $t = 100$.

que otra vez muestra que la distancia entre los bloques unitarios A_{10} y A_{11} (La Matanza y CABA) es pequeña, pero, a diferencia de $d_{100,H}$, A_{10} y A_{11} no son las regiones más cercanas en d_{100}^{WH} . Notar que $d_{100}^{WH}(6, 7) < d_{100}^{WH}(9, 10)$, 9 y 10 en la Figura 4.4.6 son la fila y columna que representan a A_{10} y A_{11} , puesto que en el formato de Python enumera de 0 a 15. Por otro lado 6 y 7 corresponde a los conglomerados {Berazategui, Florencio Varela, Presidente Perón, Almirante Brown} y {Lanús, Avellaneda, Quilmes}. En estas consideraciones generales conviene recordar la no linealidad de los algoritmos involucrados en el cálculo espectral.

4.5. Casi-métrica de Hausdorff

El teorema de metrización de espacios casi métricos, ver [MS79] y [Gus74] permite extender la noción de distancia de Hausdorff a espacios casi-métricos usando los resultados de la Sección 4.1. Ver también [ACI10].

PROPOSICIÓN 4.5.1. *Sea (X, ρ) un espacio casi-métrico. Sea \mathcal{K} , como antes, la familia de todos los subconjuntos compactos de X . Denotamos ρ_H con la función*

$$\rho_H(K^1, K^2) = \inf\{\epsilon > 0 : K^1 \subseteq K_\epsilon^2 \text{ y } K^2 \subseteq K_\epsilon^1\}$$

con $K_\epsilon = \bigcup_{x \in K} B_\delta(x, \epsilon)$ y $K^1, K^2 \in \mathcal{K}$. Entonces

- (a) ρ_H esta bien definida en $\mathcal{K} \times \mathcal{K}$;

- (b) ρ_H es una casi-métrica en \mathcal{K} con una constante triangular que depende únicamente de la casi-métrica ρ en X ;
- (c) existe una constante $\beta > 0$ tal que ρ_H es cuantitativamente equivalente a d_H^β siendo d_H la métrica de Hausdorff proporcionada por la Proposición 4.1.2 y d alguna métrica en X tal que ρ es cuantitativamente equivalentemente a d^β .

Usando de nuevo el teorema de metrización de espacios casi-métricos, vemos que el resultado del ítem (c) de la Proposición 4.1.2 de la Sección 4.1 se extiende al caso casi-métrico.

Dado un espacio casi-métrico cualquiera (Y, δ) usaremos la notación $\mathcal{K}(Y, \delta)$ para denotar a todos los δ -compactos de Y .

TEOREMA 4.5.2. *Sea (X, ρ) espacio casi-métrico completo entonces $(\mathcal{K}(X, \rho), \rho_H)$ también es completo*

DEMOSTRACIÓN. Por el teorema de metrización existe una métrica d y un número positivo β tales que $\rho \simeq d^\beta$. Más precisamente, existen $0 < c_1 \leq c_2 < \infty$ tales que

$$c_1 \rho(x, y) \leq d^\beta(x, y) \leq c_2 \rho(x, y)$$

para toda elección de $x, y \in X$. Es claro de esta propiedad que los ρ -abiertos coinciden con los d -abiertos. Como la compacidad es una propiedad topológica, tenemos que los compactos de X con la métrica d son los mismos que los compactos en X con la casi-métrica ρ . Por otra parte, por (c) de la Proposición 4.5.1, tenemos que la sucesión de compactos $\{K^n : n \in \mathbb{N}\}$ es de Cauchy en ρ_H si y sólo si es también de Cauchy en d_H . Por consiguiente

$$K^n \xrightarrow[d_H]{n \rightarrow \infty} K \in \mathcal{K}(X, d) = \mathcal{K}(X, \rho).$$

Como $d_H^\beta(K^n, K) \cong \rho_H(K^n, K)$, se tiene que $\rho_H(K^n, K)$ también tiende a cero cuando n tiende a infinito. En otras palabras $(\mathcal{K}(X, \rho), \rho_H)$ es completo. \square

4.6. Afinidades transitivas y casi-métricas de Hausdorff

En esta sección, haremos uso del resultado en [AG18] sobre casi-metrización de afinidades transitivas para probar un resultado de casi-metrización de afinidades de Hausdorff generadas por afinidades transitivas. El resultado está contenido en el siguiente enunciado.

TEOREMA 4.6.1. *Sea W una afinidad en X que es simétrica y que cumple la siguiente propiedad de transitividad: existe una función $\nu : \mathbb{R}^+ \rightarrow \mathbb{R}^+$, cóncava, continua, no negativa, creciente tal que $\nu(t) < t$ para todo $t > 0$ y cada vez que $W(x, y) > \lambda$ y $W(y, z) > \lambda$ resulta que $W(x, z) > \nu(\lambda)$. Entonces, existe φ decreciente y continua definida en \mathbb{R}^+ tal que*

- (a) $W(x, y) = \varphi(d(x, y))$ con d una casi-métrica en X ;
- (b) $W_H(A, B) = \varphi(d_H(A, B))$ para A, B compactos en X , donde W_H denota la afinidad de Hausdorff de la Sección 4.3 y d_H la casi-métrica de Hausdorff de la Sección 4.5.

DEMOSTRACIÓN. La parte a) esta en [AG18].

Para demostrar b), notemos que

$$\begin{aligned} B_d(x, \epsilon) &= \{y : d(x, y) < \epsilon\} \\ &= \{y : \phi(d(x, y)) > \phi(\epsilon)\} \\ &= \{y : W(x, y) > \phi(\epsilon)\} \\ &= E_W(x, \phi(\epsilon)). \end{aligned}$$

Por lo tanto $A_\epsilon = \bigcup_{x \in A} B_d(x, \epsilon) = \bigcup_{x \in A} E_W(x, \phi(\epsilon)) = E_W(A, \phi(\epsilon))$. Entonces

$$\begin{aligned} \phi(d_H(A, B)) &= \phi(\inf\{\epsilon > 0 : A \subset B_\epsilon \text{ y } B \subset A_\epsilon\}) \\ &= \sup\{\phi(\epsilon) : A \subset B_\epsilon \text{ y } B \subset A_\epsilon\} \\ &= \sup\{\phi(\epsilon) : A \subset E_W(B, \phi(\epsilon)) \text{ y } B \subset E_W(A, \phi(\epsilon))\} \\ &= W_H(A, B). \end{aligned}$$

□

Finalmente ilustramos el último resultado en el caso más simple de la afinidad que se obtiene como reciproca de la distancia en algún subconjunto finito de \mathbb{R} . Precisamente, para $X \subset \mathbb{R}$ con $\#(X) < \infty$, tomamos $\nu(\lambda) = \frac{\lambda}{2}$ y $W(x, y) = \frac{1}{|x-y|}$, tenemos que $\varphi(t) = \frac{1}{t}$. Si tomamos $X = \{\frac{i}{100} : i = 0, \dots, 99\}$ y una partición de X asociada al orden natural de \mathbb{R} en diez subconjuntos disjuntos $A_j = \{\frac{i}{100} : 10j \leq i < 10(j+1)\}$ $j = 0, 1, \dots, 9$. Tenemos, usando el algoritmo para el cálculo de afinidad de Hausdorff de la Sección 4.4,

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| 0 | 0 | 10 | 5 | 3.333333254 | 2.5 | 2 | 1.666666627 | 1.428571463 | 1.25 | 1.111111164 |
| 1 | 10 | 0 | 10 | 5 | 3.333333254 | 2.5 | 2 | 1.666666627 | 1.428571463 | 1.25 |
| 2 | 5 | 10 | 0 | 10 | 5 | 3.333333254 | 2.5 | 2 | 1.666666627 | 1.428571463 |
| 3 | 3.333333254 | 5 | 10 | 0 | 10 | 5 | 3.333333254 | 2.5 | 2 | 1.666666627 |
| 4 | 2.5 | 3.333333254 | 5 | 10 | 0 | 10 | 5 | 3.333333254 | 2.5 | 2 |
| 5 | 2 | 2.5 | 3.333333254 | 5 | 10 | 0 | 10 | 5 | 3.333333254 | 2.5 |
| 6 | 1.666666627 | 2 | 2.5 | 3.333333254 | 5 | 10 | 0 | 10 | 5 | 3.333333254 |
| 7 | 1.428571463 | 1.666666627 | 2 | 2.5 | 3.333333254 | 5 | 10 | 0 | 10 | 5 |
| 8 | 1.25 | 1.428571463 | 1.666666627 | 2 | 2.5 | 3.333333254 | 5 | 10 | 0 | 10 |
| 9 | 1.111111164 | 1.25 | 1.428571463 | 1.666666627 | 2 | 2.5 | 3.333333254 | 5 | 10 | 0 |

FIGURA 4.6.7. Matriz W_H .

“en cambio”, emulando a Borges en “Pierre Menard, autor del Quijote” [Bor44], usando el algoritmo para el cálculo de la distancia de Hausdorff de la Sección 4.2 y componiendo con $\varphi(t) = \frac{1}{t}$, tenemos

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| 0 | inf | 10 | 5 | 3.333333254 | 2.5 | 2 | 1.666666627 | 1.428571463 | 1.25 | 1.111111164 |
| 1 | 10 | inf | 10 | 5 | 3.333333254 | 2.5 | 2 | 1.666666627 | 1.428571463 | 1.25 |
| 2 | 5 | 10 | inf | 10 | 5 | 3.333333254 | 2.5 | 2 | 1.666666627 | 1.428571463 |
| 3 | 3.333333254 | 5 | 10 | inf | 10 | 5 | 3.333333254 | 2.5 | 2 | 1.666666627 |
| 4 | 2.5 | 3.333333254 | 5 | 10 | inf | 10 | 5 | 3.333333254 | 2.5 | 2 |
| 5 | 2 | 2.5 | 3.333333254 | 5 | 10 | inf | 10 | 5 | 3.333333254 | 2.5 |
| 6 | 1.666666627 | 2 | 2.5 | 3.333333254 | 5 | 10 | inf | 10 | 5 | 3.333333254 |
| 7 | 1.428571463 | 1.666666627 | 2 | 2.5 | 3.333333254 | 5 | 10 | inf | 10 | 5 |
| 8 | 1.25 | 1.428571463 | 1.666666627 | 2 | 2.5 | 3.333333254 | 5 | 10 | inf | 10 |
| 9 | 1.111111164 | 1.25 | 1.428571463 | 1.666666627 | 2 | 2.5 | 3.333333254 | 5 | 10 | inf |

FIGURA 4.6.8. Matriz $\varphi(d_H(A, B))$.

en total correspondencia, para los índices no diagonales, con el resultado del Teorema 4.6.1.

Una distancia entre grafos métricos. El enfoque de Gromov-Lipschitz

5.1. Introducción

Puesto que un espacio métrico es un par (X, d) con X un conjunto y d una métrica en X , un espacio métrico (Y, δ) será diferente de (X, d) si alguna de sus dos componentes difieren. Es decir si $X \neq Y$ o, si aún siendo $X = Y$, se tiene que $d \neq \delta$. Por considerar ejemplos elementales tomamos primero $X = [0, 1]$ el intervalo unitario de \mathbb{R} , y $d(x_1, x_2) = |x_1 - x_2|$. Sea $Y = [0, \frac{1}{2}]$ y $\delta(y_1, y_2) = |y_1 - y_2|$. Estos dos espacios no son isométricos ya que el primero tiene el doble del diámetro del segundo. En cambio, si consideramos los espacios $X = \{(x_1, 0) \in \mathbb{R}^2 : 0 \leq x_1 \leq 1\}$ e $Y = \{(0, y_2) \in \mathbb{R}^2 : 0 \leq y_2 \leq 1\}$ con d la métrica que X hereda de \mathbb{R}^2 y δ la métrica que Y hereda de \mathbb{R}^2 , son isométricos a través de la aplicación $f : X \rightarrow Y$ $f(x, 0) = (0, x)$. Y todas sus propiedades métricas son las mismas. Esto permite, en un sentido preciso, identificar los espacios métricos que son isométricos. Formalmente, si \mathcal{M} es la familia de todos los espacios métricos, definimos la relación de equivalencia $(X, d) \sim (Y, \delta)$ si y sólo si existe una isometría entre ellos. Es decir, si existe $f : X \rightarrow Y$ tal que f es biyectiva y $\delta(f(x_1), f(x_2)) = d(x_1, x_2)$ para toda elección de x_1 y x_2 en X . Es claro que \sim define una relación de equivalencia en \mathcal{M} . Seguiremos denotando con la misma letra \mathcal{M} al espacio cociente \mathcal{M}/\sim que identifica y considera iguales a los espacios que son isométricos.

Si volvemos al ejemplo elemental precedente en el que $(X, d) = ([0, 1], |x_1 - x_2|)$, $(Y, \delta) = ([0, \frac{1}{2}], |y_1 - y_2|)$, que no definen la misma clase en \mathcal{M} , tenemos la impresión que (X, d) “está más cerca” de (Y, δ) que de $([0, \frac{1}{4}], |y_1 - y_2|)$ y “más lejos” de (Y, δ) que de $([0, \frac{3}{4}], |y_1 - y_2|)$. Aunque, claro, estas percepciones carecen de sentido ya que no tenemos, todavía, nociones de distancias en \mathcal{M} .

El objetivo de este capítulo es introducir y ejemplificar una de las métricas en \mathcal{M} que usaremos luego, en la parte central de la tesis, para metrizar, la clase de los espacios métricos con medidas y de los espacios de tipo homogéneo.

La fuente indiscutida para la construcción de metrificaciones de \mathcal{M} y de subclases de \mathcal{M} es el libro de Misha Gromov “Metric structures for Riemannian and non-Riemannian spaces” denotado [Gro99] en la bibliografía. Ver también el libro [BBI01].

5.2. La métrica de Gromov-Lipschitz

Si X e Y son dos espacios topológicos, se dice que X e Y son homeomorfos si existe entre ellos una biyección continua y con inversa continua. Los espacios topológicos cuyas topologías son inducidas por métricas admiten morfismos más cuantitativos que los homeomorfismos. Por ejemplo los espacios métricos $([0, 1], |x_1 - x_2|)$ y $([0, 1], \sqrt{|x_1 - x_2|})$ son homeomorfos ya que la topología inducida en $[0, 1]$ por $\sqrt{|x_1 - x_2|}$ es la misma que la topología inducida por $|x_1 - x_2|$. Pero sus maneras de medir distancias entre puntos del mismo conjunto $[0, 1]$ son diferentes

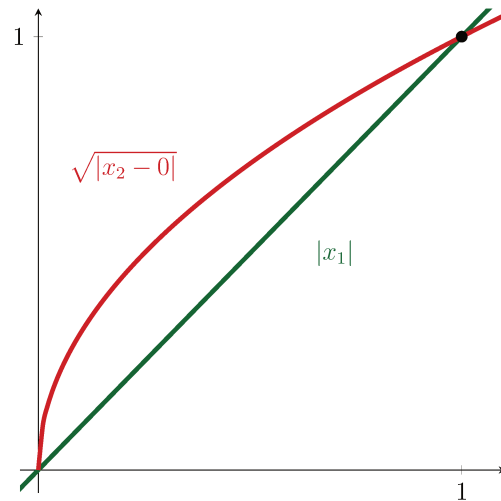


FIGURA 5.2.1. Dos métricas en $[0, 1]$.

Una pregunta natural es ¿cuán distantes están esos dos espacios métricos?. Lo mismo ocurre con ejemplos aún más elementales. Los espacios métricos $([0, 1], |x_1 - x_2|)$ y $([0, \frac{1}{2}], |y_1 - y_2|)$ son diferentes, ¿podemos cuantificar esta diferencia?.

Comenzamos en esta sección con el enfoque de Gromov-Lipschitz para definir una distancia entre espacios de una clase especial. Aquellos que son bi-Lipschitz equivalentes.

DEFINICIÓN 5.2.1.

- a) Sean $(X, d), (Y, \delta)$ dos espacios métricos. Una función $f : X \rightarrow Y$ es de **Lipschitz** si existe una constante positiva λ tal que

$$\delta(f(x_1), f(x_2)) \leq \lambda d(x_1, x_2)$$

para todo par de puntos x_1, x_2 en X .

- b) Dada $f : X \rightarrow Y$ el **coeficiente de dilatación** de f está definido por

$$\text{dil}(f) = \sup_{x_1 \neq x_2} \frac{\delta(f(x_1), f(x_2))}{d(x_1, x_2)}.$$

- c) Si $f : X \rightarrow Y$ es biyectiva, de Lipschitz y su inversa f^{-1} también es de Lipschitz, se dice que f es **bi-Lipschitz**.
- d) Una función $f : X \rightarrow Y$ es una **isometría** si es una biyección entre X e Y y además

$$d(x_1, x_2) = \delta(f(x_1), f(x_2))$$

para todos $x_1 \in X$ y $x_2 \in X$.

El siguiente enunciado reúne propiedades elementales de estos conceptos.

PROPOSICIÓN 5.2.2. Sean (X, d) e (Y, δ) dos espacios métricos

- i) $f : X \rightarrow Y$ es Lipschitz si y sólo si $\text{dil}(f) < +\infty$;
- ii) una isometría es lo mismo que una función bi-Lipschitz tal que

$$\text{dil}(f) = \text{dil}(f^{-1}) = 1;$$

- iii) las transformaciones bi-Lipschitz son homeomorfismos;
- iv) dos espacios isométricos tienen el mismo diámetro.

DEFINICIÓN 5.2.3. Sean (X, d) e (Y, δ) dos espacios métricos. Sea Λ la familia de todos los homeomorfismos bi-Lipschitz $f : X \rightarrow Y$. Cuando $\Lambda \neq \emptyset$, definimos la distancia de Gromov-Lipschitz entre (X, d) e (Y, δ) como

$$d_{GL}((X, d), (Y, \delta)) = \inf_{f \in \Lambda} (|\log \text{dil}(f)| + |\log \text{dil}(f^{-1})|)$$

y, si $\Lambda = \emptyset$,

$$d_{GL}((X, d), (Y, \delta)) = +\infty.$$

PROPOSICIÓN 5.2.4. d_{GL} tiene las siguientes propiedades

- a) $d_{GL}((X, d), (X, d)) = 0$;
- b) $d_{GL}((X, d), (Y, \delta)) = d_{GL}((Y, \delta), (X, d))$;
- c) $d_{GL}((X, d), (Y, \delta)) < \infty$ si y sólo si existe alguna transformación bi-Lipschitz entre X e Y ;
- d) $d_{GL}((X_1, d_1), (X_3, d_3)) \leq d_{GL}((X_1, d_1), (X_2, d_2)) + d_{GL}((X_2, d_2), (X_3, d_3))$;
- e) $d_{GL}((X, d), (Y, \delta)) = 0$ implica que (X, d) e (Y, δ) son isométricos, si ambos (X, d) e (Y, δ) son compactos.

DEMOSTRACIÓN. a) $f(x) = x$ es bi-Lipschitz con $\text{dil}(f) = \text{dil}(f^{-1}) = 1$. Entonces $\log \text{dil}(f) = \log \text{dil}(f^{-1}) = 0$ y $d_{GL}((X, d), (X, d)) = 0$.

b) Notar que $\Lambda(Y, X) = \{f : f^{-1} \in \Lambda(X, Y)\}$. Por consiguiente

$$\inf_{f \in \Lambda(X, Y)} (|\log \text{dil}(f)| + |\log \text{dil}(f^{-1})|) = \inf_{f \in \Lambda(Y, X)} (|\log \text{dil}(f)| + |\log \text{dil}(f^{-1})|).$$

c) Es clara por definición de d_{GL} .

d) Basta notar que si (X_i, d_i) con $i = 1, 2, 3$ son tres espacios métricos, $f \in \Lambda(X_1, X_2)$ y $g \in \Lambda(X_2, X_3)$, entonces $g \circ f \in \Lambda(X_1, X_3)$ y además $\text{dil}(g \circ f) \leq \text{dil}(g) \cdot \text{dil}(f)$ y $\text{dil}(g \circ f)^{-1} = \text{dil}(f^{-1} \circ g^{-1}) \leq \text{dil}(f^{-1}) \cdot \text{dil}(g^{-1})$. En efecto, si $x_1^1 \neq x_2^1$ son dos puntos de X_1 ,

$$\begin{aligned} \frac{d_3((g \circ f)(x_1^1), (g \circ f)(x_2^1))}{d_1(x_1^1, x_2^1)} &= \frac{d_3(g(f(x_1^1)), g(f(x_2^1)))}{d_2(f(x_1^1), f(x_2^1))} \cdot \frac{d_2(f(x_1^1), f(x_2^1))}{d_1(x_1^1, x_2^1)} \\ &\leq \text{dil}(g) \cdot \text{dil}(f). \end{aligned}$$

De manera análoga con las inversas, por consiguiente,

$$\begin{aligned} |\log \text{dil}(f \circ g)| + |\log \text{dil}(f \circ g)^{-1}| &\leq \\ &\leq |\log \text{dil}(f)| + |\log \text{dil}(g)| + |\log \text{dil}(g^{-1})| + |\log \text{dil}(f^{-1})| \\ &= (|\log \text{dil}(f)| + |\log \text{dil}(f^{-1})|) + (|\log \text{dil}(g)| + |\log \text{dil}(g^{-1})|) \\ &\leq d_{GL}((X_2, d_1), (X_2, d_2)) + d_{GL}((X_2, d_2), (X_3, d_3)) + \epsilon, \end{aligned}$$

si elegimos adecuadamente $f \in \Lambda(X_1, X_2)$ y $g \in \Lambda(X_2, X_3)$, siempre que ninguna de ellas sea vacía.

e) Para esta prueba ver [Gro99].

□

5.3. Metrización de Gromov-Lipschitz de una clase de grafos métricos

Mantendremos a lo largo de todo el resto de la tesis la notación que en esta sección introducimos. Sea (Ξ, ρ) un espacio métrico dado y fijo. En alguna circunstancia y con algunos ajustes de constantes (Ξ, ρ) podría también ser un espacio casi-métrico. Sea $n \in \mathbb{N}$ y $\mathcal{V} = \{1, 2, \dots, n\}$. También consideramos dados y fijos un número positivo R y un conjunto $\{\xi_j^0 : j = 1, \dots, n\}$ de puntos $4R$ dispersos en Ξ . Es decir $\rho(\xi_j^0, \xi_k^0) \geq 4R$ para todo $j \neq k$.

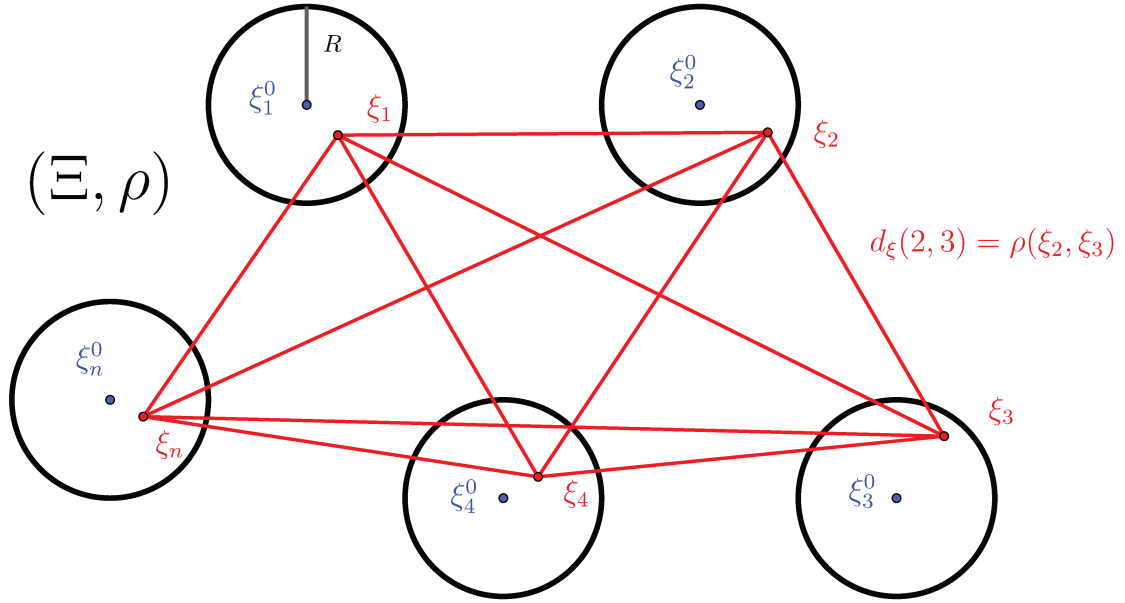


FIGURA 5.3.2. Los puntos ξ_j^0 en (Ξ, ρ) .

Sea $\Xi^n = \Xi \times \Xi \times \dots \times \Xi$ el conjunto de las n -uplas $\bar{\xi} = (\xi_1, \dots, \xi_n)$ con $\xi_j \in \Xi$, sea $\rho^n(\bar{\xi}, \bar{\eta}) = \max\{\rho(\xi_i, \eta_i) : i = 1, \dots, n\}$. El espacio (Ξ^n, ρ^n) es también métrico. Consideremos los “datos iniciales” $\{\xi_i^0 : i = 1, \dots, n\}$ como las componentes de un elemento $\bar{\xi}^0 \in \Xi^n$, es decir $\bar{\xi}^0 = (\xi_1^0, \dots, \xi_n^0)$. Sea

$$\begin{aligned} B_{\rho^n}(\bar{\xi}^0, R) &= \{\bar{\xi} \in \Xi^n : \rho^n(\bar{\xi}, \bar{\xi}^0) < R\} \\ &= \{\bar{\xi} \in \Xi^n : \rho(\xi_i, \xi_i^0) < R \text{ para todo } i = 1, \dots, n\} \\ &= \prod_{i=1}^n B_{\rho}(\xi_i^0, R), \end{aligned}$$

la bola de radio R y centro $\bar{\xi}^0$ en Ξ^n .

Construimos una familia de métricas en $\mathcal{V} = \{1, \dots, n\}$ indexadas por $\bar{\xi} \in B_{\rho^n}(\bar{\xi}^0, R)$, de la siguiente manera. Si $\bar{\xi} \in B_{\rho^n}(\bar{\xi}^0, R)$ definimos

$$d_{\bar{\xi}}(i, j) = \rho(\xi_i, \xi_j), \quad i, j \in \mathcal{V}.$$

PROPOSICIÓN 5.3.1. *Para cada $\bar{\xi} \in B_{\rho^n}(\bar{\xi}^0, R)$ se tiene $d_{\bar{\xi}}$ es una métrica en el conjunto $\mathcal{V} = \{1, \dots, n\}$.*

DEMOSTRACIÓN. Claro que $d_{\bar{\xi}}(i, i) = \rho(\xi_i, \xi_i) = 0$ para cada $i \in \mathcal{V}$. Por otra parte, si $d_{\bar{\xi}}(i, j) = 0$, resulta que $\rho(\xi_i, \xi_j) = 0$ y entonces $\xi_i = \xi_j$. Como $\xi_i \in B_{\rho}(\xi_i^0, R)$ y $\xi_j \in B_{\rho}(\xi_j^0, R)$, resulta que $B_{\rho}(\xi_i^0, R) \cap B_{\rho}(\xi_j^0, R) \neq \emptyset$. Entonces $\rho(\xi_i^0, \xi_j^0) < 2R$. Como el conjunto $\{\xi_i^0 : i \in \mathcal{V}\}$ es $4R$ disperso, esto sólo puede ocurrir si $i = j$. La simetría de cada $d_{\bar{\xi}}$ se deduce de la simetría de ρ . También la desigualdad triangular se obtiene de la desigualdad triangular de ρ ,

$$\begin{aligned} d_{\bar{\xi}}(i, k) &= \rho(\xi_i, \xi_k) \\ &\leq \rho(\xi_i, \xi_j) + \rho(\xi_j, \xi_k) \\ &= d_{\bar{\xi}}(i, j) + d_{\bar{\xi}}(j, k). \end{aligned}$$

□

Puesto que los espacios métricos $(\mathcal{V}, d_{\bar{\xi}})$ tienen el mismo cardinal y son finitos la distancia de Lipschitz-Gromov entre dos cualesquiera de ellos es finita.

Sea $\mathcal{M}(\mathcal{V}, (\Xi, \rho), \bar{\xi}^0, R) = \{(\mathcal{V}, d_{\bar{\xi}}) : \bar{\xi} \in B_{\rho^n}(\bar{\xi}^0, R)\}$. Por brevedad denotamos $\mathcal{V}_{\bar{\xi}} = (\mathcal{V}, d_{\bar{\xi}})$. Dados $\mathcal{V}_{\bar{\xi}}$ y $\mathcal{V}_{\bar{\eta}}$ con $\bar{\xi}, \bar{\eta} \in B_{\rho^n}(\bar{\xi}^0, R)$ decimos que $\mathcal{V}_{\bar{\xi}}$ es equivalente a $\mathcal{V}_{\bar{\eta}}$ y denotamos $\mathcal{V}_{\bar{\xi}} \sim \mathcal{V}_{\bar{\eta}}$ si y sólo si existe $\sigma : \mathcal{V} \rightarrow \mathcal{V}$ permutación de \mathcal{V} tal que $d_{\bar{\xi}}(\sigma(i), \sigma(j)) = d_{\bar{\eta}}(i, j)$ para $i, j \in \mathcal{V}$. En otras palabras $\mathcal{V}_{\bar{\xi}} \sim \mathcal{V}_{\bar{\eta}}$ si y sólo si $\mathcal{V}_{\bar{\xi}}$ y $\mathcal{V}_{\bar{\eta}}$ son isométricos. Entonces \sim resulta una relación de equivalencia en $\mathcal{M}(\mathcal{V}, (\Xi, \rho), \bar{\xi}^0, R)$. Sea $\widetilde{\mathcal{M}}(\mathcal{V}, (\Xi, \rho), \bar{\xi}^0, R) = \mathcal{M}(\mathcal{V}, (\Xi, \rho), \bar{\xi}^0, R) / \sim$, el conjunto cociente. Denotamos también con $\widetilde{\mathcal{V}}_{\bar{\xi}}$ a la clase de equivalencia de $\mathcal{V}_{\bar{\xi}}$. Precisamente $\widetilde{\mathcal{V}}_{\bar{\xi}} = \{\mathcal{V}_{\bar{\eta}} : \mathcal{V}_{\bar{\eta}} \sim \mathcal{V}_{\bar{\xi}}\}$.

PROPOSICIÓN 5.3.2. *Sean $\mathcal{V}_{\bar{\xi}'}$ y $\mathcal{V}_{\bar{\xi}''}$ en $\widetilde{\mathcal{V}}_{\bar{\xi}}$ y $\mathcal{V}_{\bar{\eta}'}$ y $\mathcal{V}_{\bar{\eta}''}$ en $\widetilde{\mathcal{V}}_{\bar{\eta}}$, entonces se tiene que $d_{GL}(\mathcal{V}_{\bar{\xi}'}, \mathcal{V}_{\bar{\eta}'}) = d_{GL}(\mathcal{V}_{\bar{\xi}''}, \mathcal{V}_{\bar{\eta}''})$.*

DEMOSTRACIÓN. Como $\mathcal{V}_{\bar{\xi}'} \sim \mathcal{V}_{\bar{\xi}''}$ y $\mathcal{V}_{\bar{\eta}'} \sim \mathcal{V}_{\bar{\eta}''}$, existen dos permutaciones de \mathcal{V} , σ_1 y σ_2 tales que $\sigma_1 : \mathcal{V}_{\bar{\xi}'} \rightarrow \mathcal{V}_{\bar{\xi}''}$ es isometría, es decir $d_{\bar{\xi}''}(\sigma_1(i), \sigma_1(j)) = d_{\bar{\xi}'}(i, j)$ y $\sigma_2 : \mathcal{V}_{\bar{\eta}'} \rightarrow \mathcal{V}_{\bar{\eta}''}$ también es isometría, es decir $d_{\bar{\eta}''}(\sigma_2(i), \sigma_2(j)) = d_{\bar{\eta}'}(i, j)$. Sea ahora f una transformación bi-Lipschitz entre $\mathcal{V}_{\bar{\xi}'}$ y $\mathcal{V}_{\bar{\eta}'}$, entonces

$$\begin{array}{ccc} \mathcal{V}_{\bar{\xi}'} & \xrightarrow{f} & \mathcal{V}_{\bar{\eta}'} \\ \sigma_1 \downarrow & & \downarrow \sigma_2 \\ \mathcal{V}_{\bar{\xi}''} & \xrightarrow{\sigma_2 \circ f \circ \sigma_1^{-1}} & \mathcal{V}_{\bar{\eta}''} \end{array}$$

calculemos $\text{dil}(\sigma_2 \circ f \circ \sigma_1^{-1})$,

$$\begin{aligned} \frac{d_{\bar{\eta}''}(\sigma_2 \circ f \circ \sigma_1^{-1}(i), \sigma_2 \circ f \circ \sigma_1^{-1}(j))}{d_{\bar{\xi}''}(i, j)} &= \frac{d_{\bar{\eta}''}(\sigma_2(f \circ \sigma_1^{-1}(i)), \sigma_2(f \circ \sigma_1^{-1}(j)))}{d_{\bar{\xi}''}(\sigma_1(\sigma_1^{-1}(i)), \sigma_1(\sigma_1^{-1}(j)))} \\ &= \frac{d_{\bar{\eta}'}(f(\sigma_1^{-1}(i)), f(\sigma_1^{-1}(j)))}{d_{\bar{\xi}'}(\sigma_1^{-1}(i), \sigma_1^{-1}(j))}, \end{aligned}$$

como σ_1^{-1} es biyectiva en \mathcal{V} , tenemos que $\text{dil}(\sigma_2 \circ f \circ \sigma_1^{-1}) = \text{dil}(f)$. Análogamente si g es bi-Lipschitz entre $\mathcal{V}_{\bar{\xi}''}$ y $\mathcal{V}_{\bar{\eta}''}$, tenemos que $\text{dil}(\sigma_2^{-1} \circ g \circ \sigma_1) = \text{dil}(g)$. Por consiguiente $d_{GL}(\mathcal{V}_{\bar{\xi}'}, \mathcal{V}_{\bar{\eta}'}) = d_{GL}(\mathcal{V}_{\bar{\xi}''}, \mathcal{V}_{\bar{\eta}''})$. \square

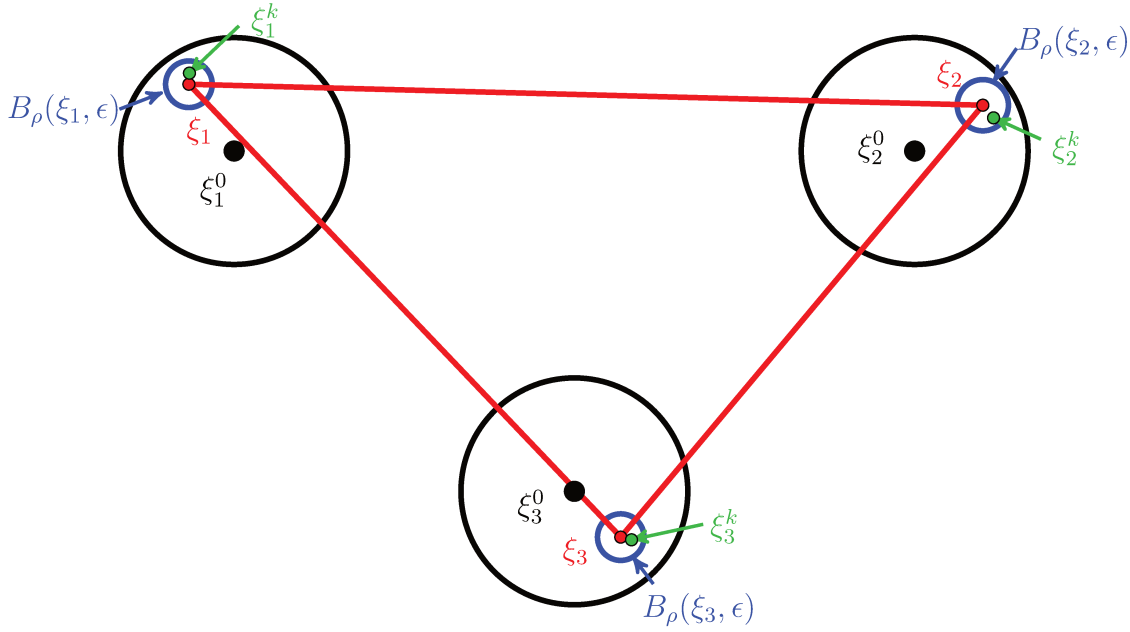
COROLARIO 5.3.3. Si definimos $\tilde{d}_{GL}(\tilde{\mathcal{V}}_{\bar{\xi}}, \tilde{\mathcal{V}}_{\bar{\eta}}) = d_{GL}(\mathcal{V}_{\bar{\xi}}, \mathcal{V}_{\bar{\eta}})$ con $\mathcal{V}_{\bar{\xi}} \in \tilde{\mathcal{V}}_{\bar{\xi}}$ y $\mathcal{V}_{\bar{\eta}} \in \tilde{\mathcal{V}}_{\bar{\eta}}$, resulta que \tilde{d}_{GL} es una métrica en $\tilde{\mathcal{M}}(\mathcal{V}, (\Xi, \rho), \bar{\xi}^0, R)$.

El siguiente resultado será útil en el Capítulo 11 cuando consideremos variables aleatorias valuadas en grafos métricos con atributos.

TEOREMA 5.3.4. Sea $\Gamma : B_{\rho^n}(\bar{\xi}^0, R) \rightarrow \tilde{\mathcal{M}}(\mathcal{V}, (\Xi, \rho), \bar{\xi}^0, R)$ dada por $\Gamma(\bar{\xi}) = \tilde{\mathcal{V}}_{\bar{\xi}}$. Entonces Γ es continua cuando en el dominio consideramos la topología generada por la métrica ρ^n y en el codominio la topología inducida por \tilde{d}_{GL} .

DEMOSTRACIÓN. Sea $\bar{\xi}^k \rightarrow \bar{\xi}$ cuando $k \rightarrow \infty$ con $\bar{\xi}^k$ y $\bar{\xi}$ todos en $B_{\rho^n}(\bar{\xi}^0, R)$, en el sentido que $\rho^n(\bar{\xi}^k, \bar{\xi}) \xrightarrow{k \rightarrow \infty} 0$. O, lo que es lo mismo, $\rho(\xi_i^k, \xi_i) \rightarrow 0$, $k \rightarrow \infty$, para todo $i \in \mathcal{V}$. Estimemos superiormente $\tilde{d}_{GL}(\tilde{\mathcal{V}}_{\bar{\xi}^k}, \tilde{\mathcal{V}}_{\bar{\xi}})$. Para ello consideramos $f : \mathcal{V}_{\bar{\xi}^k} \rightarrow \mathcal{V}_{\bar{\xi}}$ dada por $f(i) = i$, la identidad. Estimemos $\text{dil}_k(f)$. Es decir

$$\sup_{i \neq j} \frac{d_{\bar{\xi}}(i, j)}{d_{\bar{\xi}^k}(i, j)} = \sup_{i \neq j} \frac{\rho(\xi_i, \xi_j)}{\rho(\xi_i^k, \xi_j^k)}.$$

FIGURA 5.3.3. Continuidad de Γ con respecto a los parámetros.

Notar que, por la desigualdad triangular

$$\rho(\xi_i, \xi_j) \leq \rho(\xi_i, \xi_i^k) + \rho(\xi_i^k, \xi_j^k) + \rho(\xi_j^k, \xi_j) \leq 2\epsilon + \rho(\xi_i^k, \xi_j^k),$$

entonces $\rho(\xi_i, \xi_j) - 2\epsilon \leq \rho(\xi_i^k, \xi_j^k)$ y por lo tanto

$$\frac{1}{\rho(\xi_i, \xi_j) - 2\epsilon} \geq \frac{1}{\rho(\xi_i^k, \xi_j^k)}.$$

También

$$\rho(\xi_i^k, \xi_j^k) \leq \rho(\xi_i^k, \xi_i) + \rho(\xi_i, \xi_j) + \rho(\xi_j, \xi_j^k) \leq 2\epsilon + \rho(\xi_i, \xi_j),$$

entonces

$$\frac{1}{\rho(\xi_i^k, \xi_j^k)} \geq \frac{1}{\rho(\xi_i, \xi_j) + 2\epsilon}.$$

Por lo tanto

$$\frac{\rho(\xi_i, \xi_j)}{\rho(\xi_i, \xi_j) + 2\epsilon} \leq \frac{\rho(\xi_i, \xi_j)}{\rho(\xi_i^k, \xi_j^k)} \leq \frac{\rho(\xi_i, \xi_j)}{\rho(\xi_i, \xi_j) - 2\epsilon}.$$

Es decir, para todo $\epsilon > 0$, tenemos que si k es suficientemente grande, dependiendo de ϵ valen las desigualdades

$$\frac{1}{1 + 2\frac{\epsilon}{\rho(\xi_i, \xi_j)}} \leq \frac{\rho(\xi_i, \xi_j)}{\rho(\xi_i^k, \xi_j^k)} \leq \frac{1}{1 - 2\frac{\epsilon}{\rho(\xi_i, \xi_j)}},$$

para todo par $i \neq j$ en \mathcal{V} . Notemos ahora que $\rho(\xi_i, \xi_j) \geq R$ por la $4R$ dispersión del conjunto de los centros $\{\xi_1^0, \dots, \xi_n^0\}$. En efecto

$$\begin{aligned} 4R &\leq \rho(\xi_i^0, \xi_j^0) \leq \rho(\xi_i^0, \xi_i) + \rho(\xi_i, \xi_j) + \rho(\xi_j, \xi_j^0) \\ &< 2R + \rho(\xi_i, \xi_j). \end{aligned}$$

Por otra parte $\rho(\xi_i, \xi_j) \leq \sup_{i,j \in \mathcal{V}} d(\xi_i^0, \xi_j^0) + 2R = M$. Por lo tanto, para cada par $i \neq j$ y k suficientemente grande tenemos

$$\frac{1}{1 + 2\frac{\epsilon}{R}} \leq \frac{\rho(\xi_i, \xi_j)}{\rho(\xi_i^k, \xi_j^k)} \leq \frac{1}{1 - 2\frac{\epsilon}{M}}.$$

Por consiguiente

$$\frac{R}{R+S} \leq \text{dil}_k(f) \leq \frac{M}{M-S}$$

para k suficientemente grande, con $0 < R < M$ constante geométricas y $S = 2\epsilon$. Notemos que $0 < \frac{R}{R+S} < 1$ y $1 < \frac{M}{M-S} < \infty$. Como M es el diámetro de $\bigcup_{i=1}^n B_\rho(\xi_i^0, R)$, es claro que $M > R$. Para $S < M - R$ tenemos que $\frac{R+S}{R} > \frac{M}{M-S}$. En efecto

$$\begin{aligned} (R+S)(M-S) &= RM + (SM - RS - S^2) \\ &= RM + S(M - R - S) \\ &> RM. \end{aligned}$$

De aquí que

$$\begin{aligned} \left| \log \frac{R}{R+S} \right| &= \log \frac{R+S}{R} \\ &> \log \frac{M}{M-S} \\ &= \left| \log \frac{M}{M-S} \right|. \end{aligned}$$

Por consiguiente

$$\begin{aligned} |\log \text{dil}_k(f)| &\leq \log \frac{R+S}{R} \\ &= \log \left(1 + \frac{S}{R} \right) \end{aligned}$$

$$= \log \left(1 + 2 \frac{\epsilon}{R} \right).$$

Por lo tanto, dado $\epsilon > 0$, existe $k^*(\epsilon)$ tal que la desigualdad

$$d_{GL}(\mathcal{V}_{\bar{\xi}}, \mathcal{V}_{\bar{\xi}_k}) \leq \log \left(1 + 2 \frac{\epsilon}{R} \right)$$

vale para todo $k \geq k^*(\epsilon)$. Como $\log \left(1 + 2 \frac{\epsilon}{R} \right) \rightarrow 0$, $\epsilon \rightarrow 0$ tenemos que $\mathcal{V}_{\bar{\xi}_k} \xrightarrow{d_{GL}} \mathcal{V}_{\bar{\xi}}$. \square

Mientras que, a los fines de probar la medibilidad de las variables aleatorias que consideraremos luego, el teorema precedente es suficiente. En las estimaciones más cuantitativas del Capítulo 8 será importante la siguiente mejora del Teorema 5.3.4.

TEOREMA 5.3.5. *Sea Γ como en el Teorema 5.3.4. Entonces la desigualdad*

$$d_{GL}(\mathcal{V}_{\bar{\xi}}, \mathcal{V}_{\bar{\eta}}) \leq \frac{4}{R} \left(\sum_{i=1}^n \rho^2(\xi_i, \eta_i) \right)^{\frac{1}{2}}$$

vale para todo par de puntos $\bar{\xi}, \bar{\eta}$ en $B_{\rho^n}(\bar{\xi}^0, R)$ que satisfacen $\sum_{i=1}^n \rho^2(\xi_i, \eta_i) < R^2$.

DEMOSTRACIÓN. Como en el Teorema 5.3.4, para estimar superiormente $d_{GL}(\mathcal{V}_{\bar{\xi}}, \mathcal{V}_{\bar{\eta}})$, con $\bar{\xi}, \bar{\eta} \in B_{\rho^n}(\bar{\xi}^0, R)$, buscamos acotar el factor de dilatación de la aplicación bi-Lipschitz trivial dada por la identidad $f(i) = i$. Tenemos que estimar

$$\sup_{i \neq j} \frac{d_{\bar{\xi}}(i, j)}{d_{\bar{\eta}}(i, j)} = \sup_{i \neq j} \frac{\rho(\xi_i, \xi_j)}{\rho(\eta_i, \eta_j)}.$$

Notemos que $\rho(\xi_i, \xi_j) \leq [\rho(\xi_i, \eta_i) + \rho(\xi_j, \eta_j)] + \rho(\eta_i, \eta_j)$. También, para $i \neq j$ $\rho(\xi_i, \xi_j) > 2R$ y $\rho(\eta_i, \eta_j) \geq 2R$ pero $\rho(\xi_i, \eta_i) + \rho(\xi_j, \eta_j) < 2R$. Por lo tanto

$$0 < \rho(\xi_i, \xi_j) - [\rho(\xi_i, \eta_i) + \rho(\xi_j, \eta_j)] \leq \rho(\eta_i, \eta_j).$$

También

$$0 < \rho(\eta_i, \eta_j) \leq [\rho(\xi_i, \eta_i) + \rho(\xi_j, \eta_j)] + \rho(\xi_i, \xi_j).$$

La primera de las desigualdades anteriores prueba que si $S = \sup_{l \neq m} [\rho(\xi_l, \eta_l) + \rho(\xi_m, \eta_m)]$

$$0 < \rho(\xi_i, \xi_j) - S \leq \rho(\eta_i, \eta_j)$$

y la segunda prueba que

$$0 < \rho(\eta_i, \eta_j) \leq \rho(\xi_i, \xi_j) + S.$$

Ahora

$$\begin{aligned}
\frac{1}{1 + \frac{S}{R}} &\leq \frac{1}{1 + \frac{S}{\rho(\xi_i, \xi_j)}} = \frac{\rho(\xi_i, \xi_j)}{\rho(\xi_i, \xi_j) + S} \\
&\leq \frac{\rho(\xi_i, \xi_j)}{\rho(\eta_i, \eta_j)} \\
&\leq \frac{\rho(\xi_i, \xi_j)}{\rho(\xi_i, \xi_j) - S} \\
&= \frac{1}{1 - \frac{S}{\rho(\xi_i, \xi_j)}} \\
&\leq \frac{1}{1 - \frac{S}{M}},
\end{aligned}$$

donde M es el diámetro de la unión de las bolas $B(\xi^0, R)$. De aquí que el factor de dilatación de f tiene las colas

$$\frac{R}{R + S^*} \leq \text{dil}(f) \leq \frac{M}{M - S^*}$$

con $0 < R < M$ constante geométricas y $S^* = 2 \sup_l \rho(\xi_l, \eta_l)$. Notemos que $0 < \frac{R}{R+S^*} < 1$ y $1 < \frac{M}{M-S^*} < \infty$. Como M es el diámetro de $\bigcup_{i=1}^n B_\rho(\xi_i^0, R)$, es claro que $M > R$. Para $S^* < M - R$ tenemos que $\frac{R+S^*}{R} > \frac{M}{M-S^*}$. En efecto

$$\begin{aligned}
(R + S^*)(M - S^*) &= RM + (S^*M - RS^* - (S^*)^2) \\
&= RM + S^*(M - R - S^*) \\
&> RM.
\end{aligned}$$

De aquí que

$$\begin{aligned}
\left| \log \frac{R}{R + S^*} \right| &= \log \frac{R + S^*}{R} \\
&> \log \frac{M}{M - S^*} \\
&= \left| \log \frac{M}{M - S^*} \right|.
\end{aligned}$$

Por consiguiente

$$|\log \text{dil}(f)| \leq \log \frac{R + S^*}{R}$$

$$= \log \left(1 + \frac{S^*}{R} \right)$$

si $S^* \leq M - R$. Si además $S^* < R$, tenemos

$$|\log \text{dil}(f)| \leq \frac{S^*}{R} \leq \frac{2}{R} \left(\sum_{i=1}^n \rho^2(\xi_i, \eta_i) \right)^{\frac{1}{2}}.$$

□

5.4. Algunos ejemplos

En esta sección ejemplificamos, usando el lenguaje Python, algunos casos simples del cálculo de la distancia entre grafos métricos inmersos en \mathbb{R}^2 . Sea $\mathcal{G}_i = (\mathcal{V}_i, d_i)$ para $i = 1, 2$ dos grafos métricos inmersos en \mathbb{R}^2 donde d_i es la restricción de la métrica euclídea a \mathcal{V}_i . Notamos en primer lugar que, por definición de d_{GL} , resulta que $d_{GL}(\mathcal{G}_1, \mathcal{G}_2) = +\infty$ si $\#(\mathcal{V}_1) \neq \#(\mathcal{V}_2)$. Por eso nuestros ejemplos concretos estarán construidos con grafos con el mismo número de vértices y el conjunto Λ será el de todas las biyecciones entre \mathcal{V}_1 y \mathcal{V}_2 .

Describamos los pasos básicos del algoritmo. En primer lugar notamos que si $\#(\mathcal{V}_1) = \#(\mathcal{V}_2) = n$, entonces

$$\begin{aligned} \Lambda &= \{f : \mathcal{V}_1 \rightarrow \mathcal{V}_2 \text{ biyectivas}\} \\ &= \{(v_j^1, v_{\sigma(j)}^2) \mid j = 1, 2, \dots, n : \sigma \text{ permutaciones de } \{1, \dots, n\}\} \\ &= \{(v_1^1, v_{\sigma(1)}^2), (v_2^1, v_{\sigma(2)}^2), \dots, (v_n^1, v_{\sigma(n)}^2) : \sigma \text{ permutaciones de } \{1, \dots, n\}\}. \end{aligned}$$

Entonces fijada la permutación σ , calculamos $\text{dil}(f_\sigma)$ y $\text{dil}(f_\sigma^{-1})$ con f_σ la función de \mathcal{V}_1 en \mathcal{V}_2 asociada a la permutación σ . Así, considerando \mathcal{V}_1 y \mathcal{V}_2 como subconjuntos de \mathbb{R}^2 con las métricas que heredan de la euclídea en el plano, calculamos

$$\text{dil}(f_\sigma) = \max_{i \neq j} \frac{|v_{\sigma(i)}^2 - v_{\sigma(j)}^2|}{|v_i^1 - v_j^1|}$$

y también

$$\begin{aligned} \text{dil}(f_\sigma^{-1}) &= \max_{i \neq j} \frac{|v_{\sigma^{-1}(i)}^1 - v_{\sigma^{-1}(j)}^1|}{|v_i^2 - v_j^2|} \\ &= \max_{l \neq m} \frac{|v_l^1 - v_m^1|}{|v_\sigma^2(l) - v_\sigma^2(m)|} \end{aligned}$$

$$= \left(\min_{i \neq j} \frac{|v_{\sigma(i)}^2 - v_{\sigma(j)}^2|}{|v_i^1 - v_j^1|} \right)^{-1}.$$

La próxima etapa calcula logaritmos de $\text{dil}(f_\sigma)$ y $\text{dil}(f_\sigma^{-1})$, luego de tomar el valor absoluto y sumarlos, se recorre Λ usando todas las permutaciones σ y finalmente se toma el mínimo sobre todos los valores.

Comenzamos con los ejemplos tomando el caso simple en el que $\#(\mathcal{V}_1) = \#(\mathcal{V}_2) = 4$. En este contexto consideramos varias situaciones que ilustran las propiedades básicas de d_{GL} contenidas en la Proposición 5.2.4.

EJEMPLO 5.4.1. Sea $\mathcal{V}_1 = \{(0, 1), (0, 2), (2, 3), (1, 1)\} = \{v_i^1\}, i = 1, 2, 3, 4$ y $\mathcal{V}_2 = \{(2, 1), (3, 2), (\frac{3}{2}, \frac{3}{2}), (1, 3)\} = \{v_i^2\}, i = 1, 2, 3, 4$ dos conjuntos de vértices en \mathbb{R}^2 . Sea $\mathcal{G}_1 = (\mathcal{V}_1, d_1)$ el grafo de color azul en la Figura 5.4.4 y sea $\mathcal{G}_2 = (\mathcal{V}_2, d_2)$ el grafo de color rojo en la Figura 5.4.4 con d_1 y d_2 las métricas de \mathbb{R}^2 heredada por \mathcal{V}_1 y \mathcal{V}_2 .

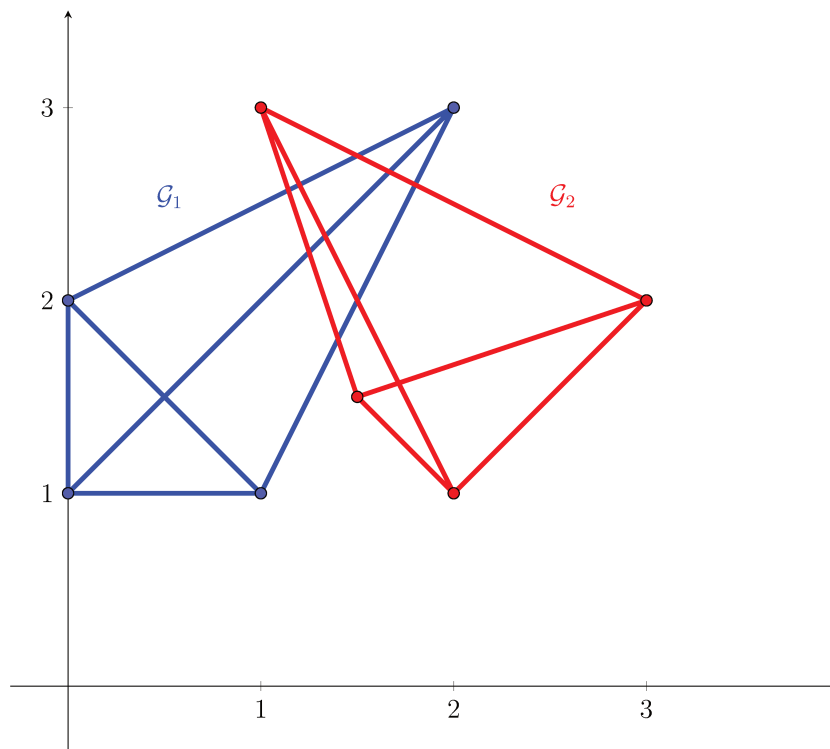


FIGURA 5.4.4. Los grafos \mathcal{G}_1 y \mathcal{G}_2 del Ejemplo 5.4.1.

Calculemos la distancia de Gromov-Lipschitz usando Python utilizando el algoritmo descrito arriba para cada σ . En la Figura 5.4.5 Se destacan los valores mínimos con color celeste que proveen $d_{GL}(\mathcal{G}_1, \mathcal{G}_2)$.

| Índice | Permutaciones f | dil(f) | log (dil(f)) | dil(inv f) | log (dil(inv f)) | log (dil(f)) + log (dil(inv f)) |
|--------|-----------------|---------|--------------|------------|------------------|-------------------------------------|
| 0 | (1, 2, 3, 4) | 2.23607 | 0.804719 | 4 | 1.38629 | 2.19101 |
| 1 | (1, 2, 4, 3) | 1.41421 | 0.346574 | 1.41421 | 0.346574 | 0.693147 |
| 2 | (1, 3, 2, 4) | 2.23607 | 0.804719 | 2 | 0.693147 | 1.49787 |
| 3 | (1, 3, 4, 2) | 1.41421 | 0.346574 | 1.41421 | 0.346574 | 0.693147 |
| 4 | (1, 4, 2, 3) | 2.23607 | 0.804719 | 2 | 0.693147 | 1.49787 |
| 5 | (1, 4, 3, 2) | 2.23607 | 0.804719 | 4 | 1.38629 | 2.19101 |
| 6 | (2, 1, 3, 4) | 2.23607 | 0.804719 | 3.16228 | 1.15129 | 1.95601 |
| 7 | (2, 1, 4, 3) | 1.58114 | 0.458145 | 2 | 0.693147 | 1.15129 |
| 8 | (2, 3, 1, 4) | 2.23607 | 0.804719 | 3.16228 | 1.15129 | 1.95601 |
| 9 | (2, 3, 4, 1) | 1.58114 | 0.458145 | 2 | 0.693147 | 1.15129 |
| 10 | (2, 4, 1, 3) | 2.23607 | 0.804719 | 3.16228 | 1.15129 | 1.95601 |
| 11 | (2, 4, 3, 1) | 2.23607 | 0.804719 | 3.16228 | 1.15129 | 1.95601 |
| 12 | (3, 1, 2, 4) | 1.58114 | 0.458145 | 1.78885 | 0.581575 | 1.03972 |
| 13 | (3, 1, 4, 2) | 1.58114 | 0.458145 | 1.78885 | 0.581575 | 1.03972 |
| 14 | (3, 2, 1, 4) | 1.58114 | 0.458145 | 4 | 1.38629 | 1.84444 |
| 15 | (3, 2, 4, 1) | 1.58114 | 0.458145 | 1.78885 | 0.581575 | 1.03972 |
| 16 | (3, 4, 1, 2) | 1.58114 | 0.458145 | 4 | 1.38629 | 1.84444 |
| 17 | (3, 4, 2, 1) | 1.58114 | 0.458145 | 1.78885 | 0.581575 | 1.03972 |
| 18 | (4, 1, 2, 3) | 2.23607 | 0.804719 | 2 | 0.693147 | 1.49787 |
| 19 | (4, 1, 3, 2) | 2.23607 | 0.804719 | 3.16228 | 1.15129 | 1.95601 |
| 20 | (4, 2, 1, 3) | 2.23607 | 0.804719 | 3.16228 | 1.15129 | 1.95601 |
| 21 | (4, 2, 3, 1) | 2.23607 | 0.804719 | 3.16228 | 1.15129 | 1.95601 |
| 22 | (4, 3, 1, 2) | 2.23607 | 0.804719 | 3.16228 | 1.15129 | 1.95601 |
| 23 | (4, 3, 2, 1) | 2.23607 | 0.804719 | 2 | 0.693147 | 1.49787 |

FIGURA 5.4.5. Aplicación del algoritmo en el Ejemplo 5.4.1, $d_{GL} \approx 0.693147$.

EJEMPLO 5.4.2. Veamos cómo reacciona Python a la propiedad e) en 5.2.4. Sea $\mathcal{V}_1 = \{(0, 0), (0, 1), (1, 1), (1, 0)\} = \{v_i^1\}, i = 1, 2, 3, 4$ el conjunto de los vértices del cuadrado $[0, 1]^2$ en \mathbb{R}^2 . Sea $\mathcal{V}_2 = \left\{ \left(3 + \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right), \left(3 + \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right), (3 + \sqrt{2}, 0), (3, 0) \right\} = \{v_i^2\}, i = 1, 2, 3, 4$. Gráficamente la situación se representa en la Figura 5.4.6 que muestra dos cuadrados del mismo lado. Claramente cualquiera de ellos se obtiene del otro componiendo traslaciones, rotaciones y permutaciones. La métrica de Gromov-Lipschitz descubre estas invariancias a partir de las relaciones métricas.

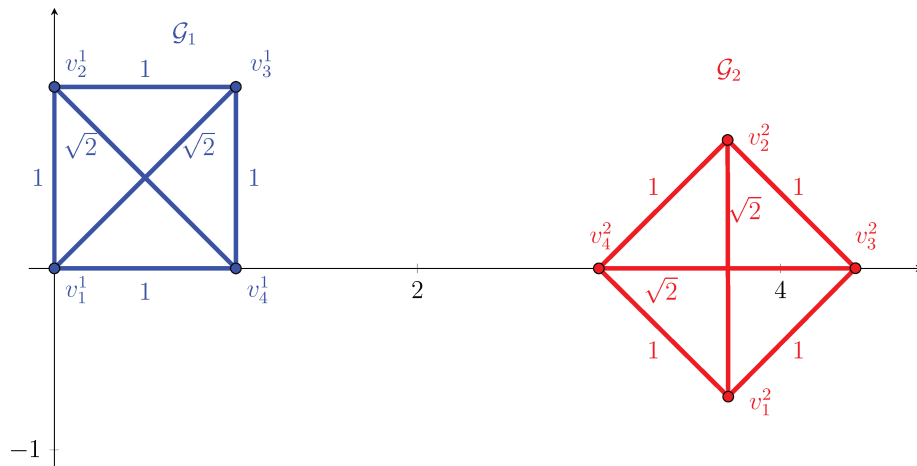


FIGURA 5.4.6. Los grafos \mathcal{G}_1 y \mathcal{G}_2 del Ejemplo 5.4.2.

En la Figura 5.4.7 se destacan los valores mínimos con color celeste, entre los correspondientes a las 24 permutaciones.

| Índice | Permutaciones f | dil(f) | $ \log(\text{dil}(f)) $ | dil(inv f) | $ \log(\text{dil}(\text{inv} f)) $ | $ \log(\text{dil}(f)) + \log(\text{dil}(\text{inv} f)) $ |
|--------|-----------------|---------|-------------------------|------------|------------------------------------|--|
| 0 | (1, 2, 3, 4) | 1.41421 | 0.346574 | 1.41421 | 0.346574 | 0.693147 |
| 1 | (1, 2, 4, 3) | 1.41421 | 0.346574 | 1.41421 | 0.346574 | 0.693147 |
| 2 | (1, 3, 2, 4) | 1 | 1.11022e-16 | 1 | 2.22045e-16 | 3.33067e-16 |
| 3 | (1, 3, 4, 2) | 1.41421 | 0.346574 | 1.41421 | 0.346574 | 0.693147 |
| 4 | (1, 4, 2, 3) | 1 | 1.11022e-16 | 1 | 2.22045e-16 | 3.33067e-16 |
| 5 | (1, 4, 3, 2) | 1.41421 | 0.346574 | 1.41421 | 0.346574 | 0.693147 |
| 6 | (2, 1, 3, 4) | 1.41421 | 0.346574 | 1.41421 | 0.346574 | 0.693147 |
| 7 | (2, 1, 4, 3) | 1.41421 | 0.346574 | 1.41421 | 0.346574 | 0.693147 |
| 8 | (2, 3, 1, 4) | 1 | 1.11022e-16 | 1 | 2.22045e-16 | 3.33067e-16 |
| 9 | (2, 3, 4, 1) | 1.41421 | 0.346574 | 1.41421 | 0.346574 | 0.693147 |
| 10 | (2, 4, 1, 3) | 1 | 1.11022e-16 | 1 | 2.22045e-16 | 3.33067e-16 |
| 11 | (2, 4, 3, 1) | 1.41421 | 0.346574 | 1.41421 | 0.346574 | 0.693147 |
| 12 | (3, 1, 2, 4) | 1.41421 | 0.346574 | 1.41421 | 0.346574 | 0.693147 |
| 13 | (3, 1, 4, 2) | 1 | 1.11022e-16 | 1 | 2.22045e-16 | 3.33067e-16 |
| 14 | (3, 2, 1, 4) | 1.41421 | 0.346574 | 1.41421 | 0.346574 | 0.693147 |
| 15 | (3, 2, 4, 1) | 1 | 1.11022e-16 | 1 | 2.22045e-16 | 3.33067e-16 |
| 16 | (3, 4, 1, 2) | 1.41421 | 0.346574 | 1.41421 | 0.346574 | 0.693147 |
| 17 | (3, 4, 2, 1) | 1.41421 | 0.346574 | 1.41421 | 0.346574 | 0.693147 |
| 18 | (4, 1, 2, 3) | 1.41421 | 0.346574 | 1.41421 | 0.346574 | 0.693147 |
| 19 | (4, 1, 3, 2) | 1 | 1.11022e-16 | 1 | 2.22045e-16 | 3.33067e-16 |
| 20 | (4, 2, 1, 3) | 1.41421 | 0.346574 | 1.41421 | 0.346574 | 0.693147 |
| 21 | (4, 2, 3, 1) | 1 | 1.11022e-16 | 1 | 2.22045e-16 | 3.33067e-16 |
| 22 | (4, 3, 1, 2) | 1.41421 | 0.346574 | 1.41421 | 0.346574 | 0.693147 |
| 23 | (4, 3, 2, 1) | 1.41421 | 0.346574 | 1.41421 | 0.346574 | 0.693147 |

FIGURA 5.4.7. Aplicación del algoritmo, $d_{GL} = 0$.

Así $d_{GL}(\mathcal{G}_1, \mathcal{G}_2) = 3.33067 \cdot 10^{-16} \simeq 0$ que corresponde a e de la Proposición 5.2.4.

EJEMPLO 5.4.3. En este ejemplo consideraremos el problema de aproximación de un grafo dado por otros. Tomamos $\mathcal{G}_1 = (\mathcal{V}_1, d_1)$ el mismo del Ejemplo 5.4.2. Es decir $v_1^1 = (0, 0), v_2^1 = (0, 1), v_3^1 = (1, 1)$ y $v_4^1 = (1, 0)$ con d_1 la distancia que \mathcal{V}_1 hereda de \mathbb{R}^2 . En cambio \mathcal{V}_2 será elegido aleatoriamente con el siguiente criterio, $\mathcal{V}_2^r(\omega) = (v_1^2(\omega, r), v_2^2(\omega, r), v_3^2(\omega, r), v_4^2(\omega, r))$, donde $v_i^2(\omega)$ es un punto de \mathbb{R}^2 elegido al azar (uniforme) en el cuadrado con centro en v_i^1 y lado $0 < r < \frac{1}{2}$. En este caso el algoritmo básico consiste en componer el algoritmo de los ejemplos anteriores, luego de seleccionar \mathcal{V}_2^r al azar de la manera descripta arriba.

En este caso podemos representar la dependencia de $d_{GL}(\mathcal{V}_1, \mathcal{V}_2^r(\omega))$ como función del grado de aproximación r en la Figura 5.4.8 que tiene comportamiento similar en distintas escalas.

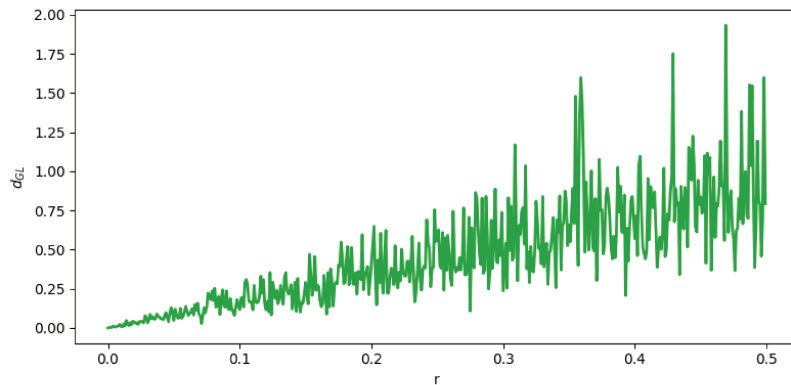


FIGURA 5.4.8. Grafica $d_{GL}(\mathcal{V}_1, \mathcal{V}_2^r(\omega))$ en función de $0 < r < 0.5$.

La Figura 5.4.9 muestra el comportamiento similar a la Figura 5.4.8 ahora en el intervalo $0 < r < 0.1$.

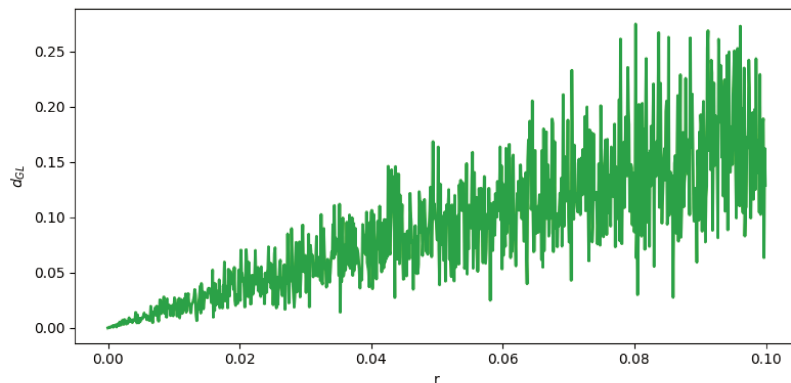


FIGURA 5.4.9. Grafica $d_{GL}(\mathcal{V}_1, \mathcal{V}_2^r(\omega))$ en función de $0 < r < 0.1$.

Cabe observar que al menos empíricamente $d_{GL}(\mathcal{V}_1, \mathcal{V}_2^r(\omega)) \leq cr$ para una constante c positiva. Más adelante en esta tesis consideraremos el problema de los grafos aleatorios, la selección aleatoria en el ejemplo precedente solo pretende ilustrar la aproximación de la d_{GL} cuando los vértices de los grafos son cercanos.

5.5. Distancia de Gromov-Lipschitz entre métricas difusivas en un grafo.

Aplicación al grafo SUBE en AMBA

Ya observamos que la igualdad de los cardinales de dos espacios métricos es necesaria para la finitud de d_{GL} . Un caso de interés práctico de tal situación, como ya vimos en la Sección 5.3, se tiene cuando $\mathcal{V}_1 = \mathcal{V}_2 = \mathcal{V}$ y en \mathcal{V} hay distintas métricas. Por ejemplo en el caso en que \mathcal{V} es el conjunto de los 41 distritos de AMBA tenemos una familia grande de métricas, las difusivas $\{d_t : t > 0\}$. En general, dado un grafo $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ no dirigido y ponderado, con atributos unitarios, tenemos definidas

$$d_t^2(l, m) = \sum_{j=1}^{n-1} e^{2t\lambda_j} |\phi_j(l) - \phi_j(m)|^2.$$

Entonces tenemos también toda una familia de espacios métricos $\{(\mathcal{V}, d_t) : t > 0\}$ sobre el mismo conjunto \mathcal{V} . Por consiguiente, podemos preguntarnos cuál es la distancia de Gromov-Lipschitz entre dos de éstos. En particular podemos preguntarnos por la continuidad de esta familia. Es decir dado $t_0 > 0$ ¿es $d_{GL}((\mathcal{V}, d_{t_0}), (\mathcal{V}, d_t)) < \epsilon$ si $|t - t_0| < \delta$? El siguiente resultado muestra que la situación es aún mejor y podemos ver que la aplicación $E : \mathbb{R}^+ \rightarrow \mathcal{M}$ (espacios métricos) dada por $E(t) = (\mathcal{V}, d_t)$ es localmente Lipschitz-1 con la métrica usual en el dominio \mathbb{R} y la de Gromov-Lipschitz en el codominio.

TEOREMA 5.5.1. *Con la notación precedente, para todo $t_0 > 0$ existen $\epsilon > 0$ y $C > 0$ tales que*

$$d_{GL}(E(t), E(t_0)) \leq C|t - t_0|$$

para todo t tal que $|t - t_0| < \epsilon$.

DEMOSTRACIÓN. Basta considerar sólo una transformación bi-Lipschitz entre (\mathcal{V}, d_{t_0}) y (\mathcal{V}, d_t) , la identidad I . Puesto que $\text{dil}(I) = \sup_{l \neq m} \frac{d_t(l, m)}{d_{t_0}(l, m)}$ y $\text{dil}(I^{-1}) = \sup_{l \neq m} \frac{d_{t_0}(l, m)}{d_t(l, m)}$, es suficiente estimar superior, inferior y uniformemente, el cociente $\frac{d_t(l, m)}{d_{t_0}(l, m)}$. Veamos esto,

con $C_1 = (\min_{l \neq m} d_{t_0}^2(l, m))^{-1}$, tenemos

$$\begin{aligned}
\left| \left(\frac{d_t(l, m)}{d_{t_0}(l, m)} \right)^2 - 1 \right| &= \frac{|d_t^2(l, m) - d_{t_0}^2(l, m)|}{d_{t_0}^2(l, m)} \\
&\leq C_1 |d_t^2(l, m) - d_{t_0}^2(l, m)| \\
&= C_1 \left| \sum_{j=1}^{n-1} e^{2t\lambda_j} |\phi_j(l) - \phi_j(m)|^2 - \sum_{j=1}^{n-1} e^{2t_0\lambda_j} |\phi_j(l) - \phi_j(m)|^2 \right| \\
&\leq C_1 \sum_{j=1}^{n-1} |e^{2t\lambda_j} - e^{2t_0\lambda_j}| |\phi_j(l) - \phi_j(m)|^2 \\
&\leq \left(C_2 \sum_{j=1}^{n-1} |\phi_j(l) - \phi_j(m)|^2 \right) |t - t_0| \\
&\leq C_3 |t - t_0|
\end{aligned}$$

si $|t - t_0|$ es suficientemente chico. Por consiguiente y achicando todavía, de ser necesario $|t - t_0|$, tenemos

$$\sqrt{1 - C_3|t - t_0|} \leq \frac{d_t(l, m)}{d_{t_0}(l, m)} \leq \sqrt{1 + C_3|t - t_0|}.$$

Entonces

$$\text{dil}(I) \leq \sqrt{1 + C_3|t - t_0|}$$

y

$$\text{dil}(I^{-1}) \leq (\sqrt{1 - C_3|t - t_0|})^{-1}.$$

Por consiguiente $\log \text{dil}(I) \leq \log \sqrt{1 + C_3|t - t_0|}$. Si $|t - t_0|$ es suficientemente chico. Además $\log \text{dil}(I^{-1}) \leq -\log \sqrt{1 - C_3|t - t_0|}$. Ambas estimaciones prueban que para $|t - t_0|$ chico,

$$|\log \text{dil}(I)| + |\log \text{dil}(I^{-1})| \leq C_4|t - t_0|.$$

Por consiguiente $d_{GL}((\mathcal{V}, d_t), (\mathcal{V}, d_{t_0})) \leq C_4|t - t_0|$ en un entorno de t_0 . \square

En principio el análisis en el teorema anterior de sólo una aplicación bi-Lipschitz, la identidad, podría hacernos pensar que la estimación obtenida es muy rudimentaria.

Sin embargo la siguiente aplicación en AMBA, muestra que el carácter Lipschitz-1 de la función $E(t)$ parece óptima.

Sea, entonces, $\mathcal{V} = \{1, \dots, 41\}$ el conjunto de vértices del grafo de SUBE en AMBA con W la matriz de la Sección 1.5 en la Figura 1.5.7 y con atributos uniformes. El algoritmo introducido en el Capítulo 1 nos provee los autovalores y las autofunciones. Los resultados de la Sección 4.2 (Ejemplo 4.2.2), nos proveen también el algoritmo para el cálculo de d_t . Tomando $t_0 = 100$, $50 \leq t \leq 200$ y calculando $d_{GL}((\mathcal{V}, d_{100}), (\mathcal{V}, d_t))$ para t en ese intervalo y con paso temporal igual a uno, se obtiene la siguiente gráfica en la Figura 5.5.10 para la función $d_{GL}((\mathcal{V}, d_{100}), (\mathcal{V}, d_t))$.

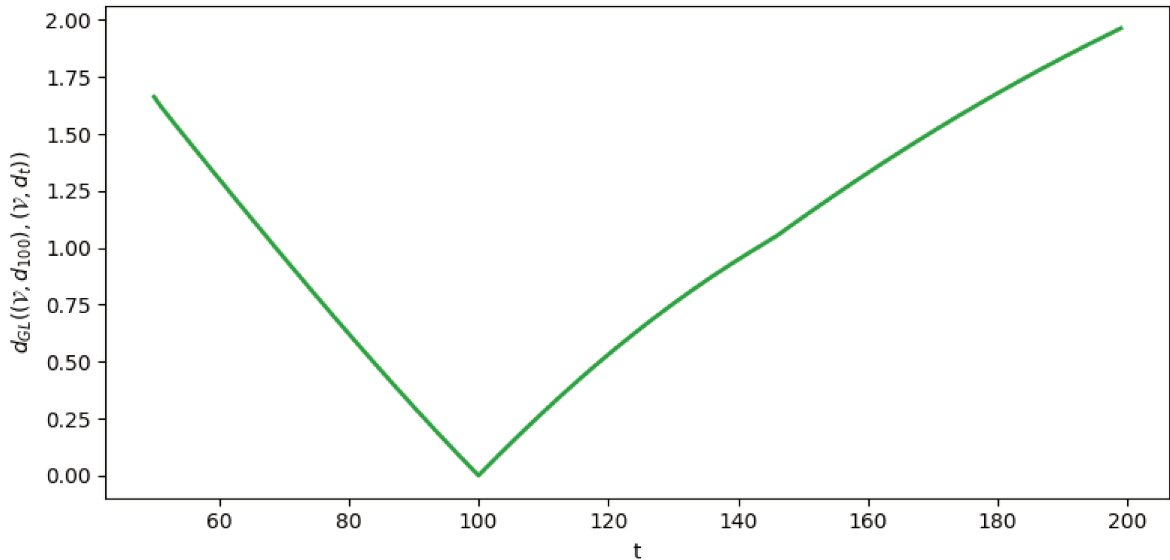


FIGURA 5.5.10. El carácter Lipschitz de la función $E(t)$ para t cercano a $t_0 = 100$.

Distancia entre grafos métricos. El enfoque Gromov-Hausdorff

6.1. Introducción

La métrica introducida en el capítulo precedente requiere, para su finitud, que los dos espacios métricos tengan el mismo cardinal, dado el carácter biyectivo de las transformaciones bi-Lipschitz. Gromov propone sumergir los dos espacios métricos en análisis en un continente métrico mayor y en éste usar la distancia de Hausdorff, que expusimos en el Capítulo 4. En particular esta hace posible calcular distancias entre grafos métricos con distintos números de vértices.

6.2. La métrica de Gromov-Hausdorff

Sean (X_1, d_1) y (X_2, d_2) dos espacios métricos. Decimos que (X_i, d_i) , $i = 1, 2$, tienen una **inmersión isométrica** en el espacio métrico (Y, δ) si existen $f_i : X_i \rightarrow Y$ inyectivas tales que $d_i(x_1^i, x_2^i) = \delta(f_i(x_1^i), f_i(x_2^i))$ para toda elección de x_1^i y $x_2^i \in X_i$, $i = 1, 2$.

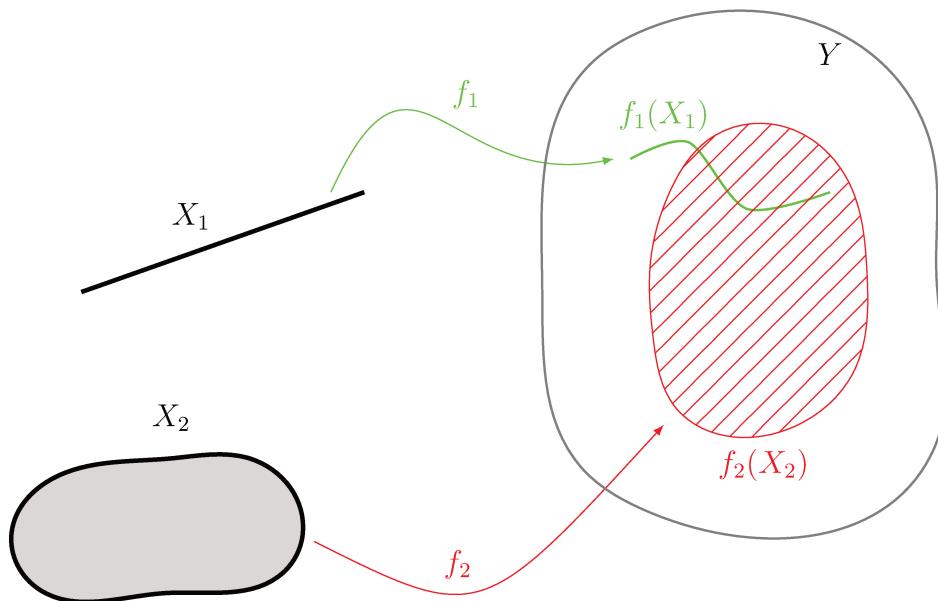


FIGURA 6.2.1. Inmersiones isométricas.

Ahora como $f_1(X_1)$ y $f_2(X_2)$ son dos subconjuntos, isométricos con X_1 y X_2 respectivamente, pero contenidos en el mismo espacio (Y, δ) podemos calcular su distancia de Hausdorff en (Y, δ) y obtener los números reales (extendidos)

$$\{d_H^{(Y, \delta)}(f_1(X_1), f_2(X_2)) : (Y, \delta) \text{ e.m. } f_i : X_i \rightarrow Y \text{ inmersiones isométricas } i = 1, 2\}.$$

El ínfimo de este conjunto, que será infinito si resulta vacío, es la **distancia de Gromov-Hausdorff** entre los espacios métricos (X_1, d_1) y (X_2, d_2) . Usamos $d_{GH}((X_1, d_1), (X_2, d_2))$ para denotar esta distancia.

Otra manera de formular la definición de d_{GH} que no alude a las nociones de isometría de inmersiones se puede dar considerando la unión disjunta de X_1 y X_2 . Dados dos conjuntos X_1 y X_2 , denotamos por $X_1 \sqcup X_2$ a su unión disjunta. Por ejemplo $X_1 \sqcup X_2 = (X_1 \times \{1\}) \cup (X_2 \times \{2\}) = \tilde{X}_1 \cup \tilde{X}_2$. Definimos d_i en \tilde{X}_i como d_i en X_i con $i = 1, 2$. Precisamente $d_1((x_1^1, 1), (x_2^1, 1)) = d_1(x_1^1, x_2^1)$ y $d_2((x_1^2, 2), (x_2^2, 2)) = d_2(x_1^2, x_2^2)$. Sea $D(d_1, d_2) = \{\rho : \text{métrica en } X_1 \sqcup X_2 \text{ con } \rho|_{X_1} = d_1 \text{ y } \rho|_{X_2} = d_2\}$. Con esta notación tenemos el siguiente resultado cuya demostración puede hallarse en [BBI01].

PROPOSICIÓN 6.2.1. Sean (X_1, d_1) y (X_2, d_2) dos espacios métricos. Entonces $d_{GH}((X_1, d_1), (X_2, d_2)) = \inf\{d_{H, \rho}(\tilde{X}_1, \tilde{X}_2), \text{ donde } d_{H, \rho} \text{ es la distancia de Hausdorff en } (X_1 \sqcup X_2, \rho), \rho \in D(d_1, d_2)\}$.

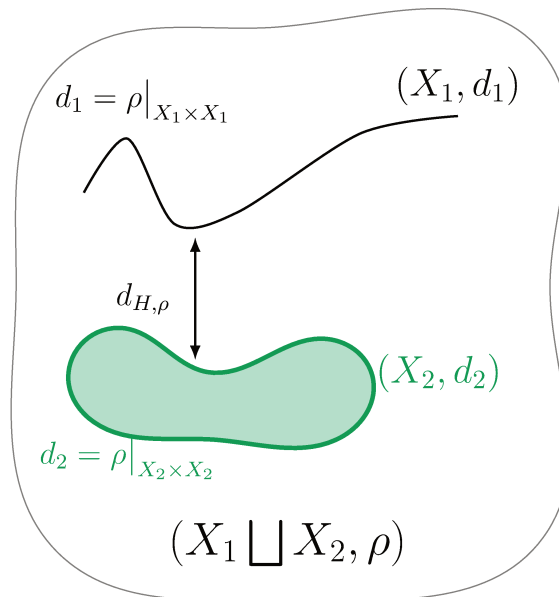


FIGURA 6.2.2. Esquema de la Proposición 6.2.1.

PROPOSICIÓN 6.2.2. *La función $d_{GH}((X_1, d_1), (X_2, d_2))$ satisface las siguientes propiedades.*

- a) $d_{GH}((X, d), (X, d)) = 0$;
- b) si (X_1, d_1) y (X_2, d_2) son isométricos entonces $d_{GH}((X_1, d_1), (X_2, d_2)) = 0$;
- c) $d_{GH}((X_1, d_1), (X_2, d_2)) = d_{GH}((X_2, d_2), (X_1, d_1))$ para todo par de espacios métricos (X_i, d_i) , $i = 1, 2$;
- d) $d_{GH}((X_1, d_1), (X_3, d_3)) \leq d_{GH}((X_1, d_1), (X_2, d_2)) + d_{GH}((X_2, d_2), (X_3, d_3))$ para toda terna de espacios métricos (X_i, d_i) , $i = 1, 2, 3$;
- e) si (X_i, d_i) , $i = 1, 2$ son compactos y si $d_{GH}((X_1, d_1), (X_2, d_2)) = 0$, entonces (X_1, d_1) y (X_2, d_2) son isométricos.

Un ejemplo particular se tiene cuando queremos calcular la distancia de un subespacio métrico a otro espacio métrico que lo contiene. En este caso tenemos $X_1 \subset X_2$ y d_1 la restricción de d_2 a X_1 . Por consiguiente la inclusión f_2 de X_1 en X_2 es inmersión isométrica y la identidad f_2 de X_2 en si mismo también. Entonces

$$d_{GH}((X_1, d_1), (X_2, d_2)) \leq d_{H, d_2}(X_1, X_2).$$

Observamos que la dificultad computacional que en la definición de d_{GH} se refleja en el cardinal del conjunto de todas las posibles inmersiones isométricas f_i , en el resultado de la Proposición 6.2.1 se refleja en la diversidad de métricas en $X_1 \sqcup X_2$ que extienden a las d_i . Sin embargo resulta muchas veces sencillo obtener cotas superiores para la distancia de Gromov-Hausdorff entre dos espacio métricos, cuando alguna inmersión resulta evidente.

Consideremos por ejemplo $(X_1, d_1) = ([0, 1], |x_1 - x_2|)$ y $(X_2, d_2) = ([0, 1], \sqrt{|x_1 - x_2|})$. Probemos primero que $d_{GL} \left(([0, 1], |x_1 - x_2|), ([0, 1], \sqrt{|x_1 - x_2|}) \right) = +\infty$. En efecto si $f : (X_1, d_1) \rightarrow (X_2, d_2)$ es bi-Lipschitz tenemos que

$$c_1 \leq \frac{\sqrt{|f(x_1) - f(x_2)|}}{|x_1 - x_2|} = \sqrt{\frac{|f(x_1) - f(x_2)|}{|x_1 - x_2|}} \cdot \frac{1}{\sqrt{|x_1 - x_2|}} \leq c_2,$$

entonces

$$|x_1 - x_2|c_1^2 \leq \left| \frac{f(x_1) - f(x_2)}{x_1 - x_2} \right| \leq c_2^2|x_1 - x_2|,$$

para toda elección de x_1 y x_2 . Por consiguiente $f' \equiv 0$ en $(0, 1)$ y f sería constante.

Pero $d_{GH} \left(([0, 1], |x_1 - x_2|), ([0, 1], \sqrt{|x_1 - x_2|}) \right) < \infty$. En efecto, considerar $(Y, \delta) = (\mathbb{R}^2, \delta(x_1, y_1), (x_2, y_2))$ con $\delta(x_1, x_2) = \max \left\{ |x_1 - x_2|, \sqrt{|y_1 - y_2|} \right\}$ el plano parabólico. Entonces $f_1 : [0, 1] \rightarrow \mathbb{R}^2 ; f_1(x) = (x, 0)$ y $f_2 : [0, 1] \rightarrow \mathbb{R}^2 ; f_2(y) = (0, y)$ son dos immersiones isométricas en (Y, δ) . En efecto

$$\begin{aligned} \delta(f_1(x_1), f_1(x_2)) &= \delta((x_1, 0), (x_2, 0)) \\ &= |x_1 - x_2| \\ &= d_1(x_1, x_2), \end{aligned}$$

y

$$\begin{aligned} \delta(f_2(y_1), f_2(y_2)) &= \delta((0, y_1), (0, y_2)) \\ &= \sqrt{|y_1 - y_2|} \\ &= d_2(y_1, y_2). \end{aligned}$$

Por otra parte como $[0, 1] \times \{0\}$ y $\{0\} \times [0, 1]$ son dos compactos en el espacio métrico (\mathbb{R}^2, δ) tenemos que

$$d_{H,\delta}([0, 1] \times \{0\}, (\{0\} \times [0, 1])) < \infty.$$

Pero además puede estimarse por arriba por 1.

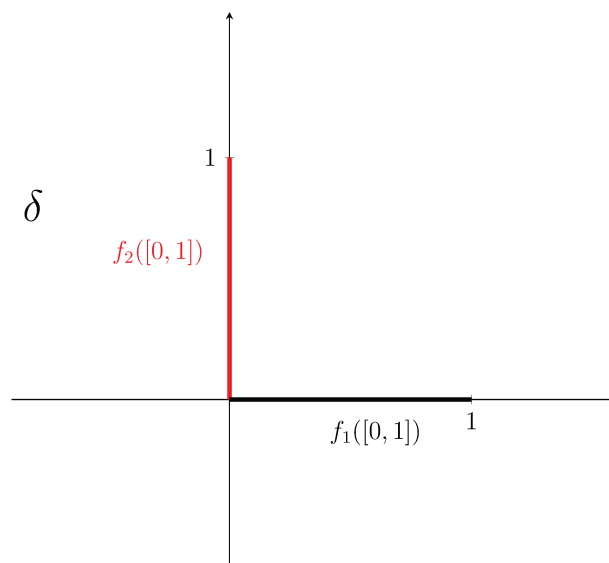


FIGURA 6.2.3. $d_{H,\delta}(f_1[0, 1], f_2[0, 1]) \leq 1$.

6.3. Metrización de Gromov-Hausdorff de una clase de grafos métricos

En esta sección volvemos a considerar la situación de la Sección 5.3 del Capítulo 5, pero ahora con la métrica de Gromov-Hausdorff en vez de la de Gromov-Lipschitz. Aunque este contexto nos permite estimar las distancias entre espacios métricos de diferentes cardinales, seguiremos con la misma situación de la Sección 5.3. Sea $n \in \mathbb{N}$ y $\mathcal{V} = \{1, 2, \dots, n\}$. Sea (Ξ, ρ) un espacio métrico dado, $R > 0$ y $\{\xi_j^0 : j = 1, \dots, n\}$ un conjunto $4R$ disperso con respecto a ρ en Ξ . Dado $\bar{\xi} \in B_{\rho^n}(\bar{\xi}^0, R)$ con $\bar{\xi}^0 = (\xi_1^0, \dots, \xi_n^0)$, tenemos la métrica $d_{\bar{\xi}}(i, j) = \rho(\xi_i, \xi_j)$. Y también tenemos $\widetilde{\mathcal{M}}(\mathcal{V}, (\Xi, \rho), \bar{\xi}^0, R)$. Para definir adecuadamente \widetilde{d}_{GH} en $\widetilde{\mathcal{M}}(\mathcal{V}, (\Xi, \rho), \bar{\xi}^0, R)$ probamos antes el siguiente resultado. Recordamos que $\mathcal{V}_{\bar{\xi}} = (\mathcal{V}, d_{\bar{\xi}})$.

PROPOSICIÓN 6.3.1. Sean $\mathcal{V}_{\bar{\xi}'} \sim \mathcal{V}_{\bar{\xi}''}$ y $\mathcal{V}_{\bar{\eta}'} \sim \mathcal{V}_{\bar{\eta}''}$ entonces $d_{GH}(\mathcal{V}_{\bar{\xi}'}, \mathcal{V}_{\bar{\eta}'}) = d_{GH}(\mathcal{V}_{\bar{\xi}''}, \mathcal{V}_{\bar{\eta}''})$.

DEMOSTRACIÓN. Como $\mathcal{V}_{\bar{\xi}'} \sim \mathcal{V}_{\bar{\xi}''}$ y $\mathcal{V}_{\bar{\eta}'} \sim \mathcal{V}_{\bar{\eta}''}$, existen dos permutaciones de \mathcal{V} , σ_1 y σ_2 tales que $\sigma_1 : \mathcal{V}_{\bar{\xi}'} \rightarrow \mathcal{V}_{\bar{\xi}''}$ es isometría, es decir $d_{\bar{\xi}''}(\sigma_1(i), \sigma_1(j)) = d_{\bar{\xi}'}(i, j)$ y $\sigma_2 : \mathcal{V}_{\bar{\eta}'} \rightarrow \mathcal{V}_{\bar{\eta}''}$ también es isometría, es decir $d_{\bar{\eta}''}(\sigma_2(i), \sigma_2(j)) = d_{\bar{\eta}'}(i, j)$. Sea ahora (\mathcal{Z}, ∂) un espacio métrico en el que ambos $\mathcal{V}_{\bar{\xi}'}$ y $\mathcal{V}_{\bar{\eta}'}$ se inyectan isométricamente. Sean f y g esas isometrías $f : \mathcal{V}_{\bar{\xi}'} \rightarrow (\mathcal{Z}, \partial)$ y $g : \mathcal{V}_{\bar{\eta}'} \rightarrow (\mathcal{Z}, \partial)$. Los siguientes diagramas muestran que $\mathcal{V}_{\bar{\xi}''}$ y $\mathcal{V}_{\bar{\eta}''}$ también se inyectan isométricamente en (\mathcal{Z}, ∂) .

$$\begin{array}{ccc}
 \mathcal{V}_{\bar{\xi}'} & \xrightarrow{f} & (\mathcal{Z}, \partial) \\
 \sigma_1 \downarrow & \nearrow F = f \circ \sigma_1^{-1} & \\
 \mathcal{V}_{\bar{\xi}''} & &
 \end{array}
 \qquad
 \begin{array}{ccc}
 \mathcal{V}_{\bar{\eta}'} & \xrightarrow{g} & (\mathcal{Z}, \partial) \\
 \sigma_2 \downarrow & \nearrow G = g \circ \sigma_2^{-1} & \\
 \mathcal{V}_{\bar{\eta}''} & &
 \end{array}$$

En efecto por las propiedades de σ_1 y σ_2 resulta que F y G también son isometrías y

$$\begin{aligned}
 d_H^{(\mathcal{Z}, \partial)}(F(\mathcal{V}_{\bar{\xi}''}), G(\mathcal{V}_{\bar{\eta}''})) &= d_H^{(\mathcal{Z}, \partial)}(f[\sigma_1^{-1}(\mathcal{V}_{\bar{\xi}''})], g[\sigma_2^{-1}(\mathcal{V}_{\bar{\eta}''})]) \\
 &= d_H^{(\mathcal{Z}, \partial)}(f(\mathcal{V}_{\bar{\xi}'}), g(\mathcal{V}_{\bar{\eta}'})).
 \end{aligned}$$

Por consiguiente tenemos que

$$d_{GH}(\mathcal{V}_{\bar{\xi}'}, \mathcal{V}_{\bar{\eta}'}) = d_{GH}(\mathcal{V}_{\bar{\xi}''}, \mathcal{V}_{\bar{\eta}''}).$$

□

COROLARIO 6.3.2. Si definimos $\tilde{d}_{GH}(\tilde{\mathcal{V}}_{\bar{\xi}}, \tilde{\mathcal{V}}_{\bar{\eta}}) = d_{GH}(\mathcal{V}_{\bar{\xi}}, \mathcal{V}_{\bar{\eta}})$ con $\mathcal{V}_{\bar{\xi}} \in \tilde{\mathcal{V}}_{\bar{\xi}}$ y $\mathcal{V}_{\bar{\eta}} \in \tilde{\mathcal{V}}_{\bar{\eta}}$ es una métrica en $\tilde{\mathcal{M}}(\mathcal{V}, (\Xi, \rho), \bar{\xi}^0, R)$.

TEOREMA 6.3.3. La función $\Gamma : B_{\rho^n}(\bar{\xi}^0, R) \rightarrow \tilde{\mathcal{M}}(\mathcal{V}, (\Xi, \rho), \bar{\xi}^0, R)$ dada por $\Gamma(\bar{\xi}) = \tilde{\mathcal{V}}_{\bar{\xi}}$ es continua si en el dominio tomamos la topología inducida por la métrica ρ^n y en el codominio la topología inducida por la métrica \tilde{d}_{GH} .

DEMOSTRACIÓN. Sea $\bar{\xi}^k \rightarrow \bar{\xi}$ con $\bar{\xi}^k$ y $\bar{\xi}$ en $B_{\rho^n}(\bar{\xi}^0, R)$ cuando $k \rightarrow \infty$. Sea $(\mathcal{Z}, \partial) = (\bigcup_{l=1}^n B(\xi_l^0, R), \rho)$. Sea $f_{\bar{\xi}} : \mathcal{V} \rightarrow (\mathcal{Z}, \partial)$ dada por $f_{\bar{\xi}}(i) = \xi_i$ para $i \in \mathcal{V}$, $\bar{\xi} = (\xi_1, \dots, \xi_i, \dots, \xi_n)$. Es claro que $f_{\bar{\xi}}$ es inyectiva. Además

$$\partial(f_{\bar{\xi}}(i), f_{\bar{\xi}}(j)) = \partial(\xi_i, \xi_j) = \rho(\xi_i, \xi_j) = d_{\bar{\xi}}(i, j)$$

para todo $i, j \in \mathcal{V}$. En otras palabras $f_{\bar{\xi}}$ es una isometría de \mathcal{V} en $(\bigcup_{l=1}^n B(\xi_l^0, R), \rho)$. Del mismo modo, y para cada $k \geq 1$ tenemos también, una sucesión de isometrías $f_{\bar{\xi}^k} : \mathcal{V} \rightarrow (\bigcup_{l=1}^n B(\xi_l^0, R), \rho)$ dadas por $f_{\bar{\xi}^k}(i) = \xi_i^k$. Notemos ahora que $f_{\bar{\xi}}(\mathcal{V}) = \{\xi_1, \dots, \xi_n\}$ y que $f_{\bar{\xi}^k}(\mathcal{V}) = \{\xi_1^k, \dots, \xi_n^k\}$. Por consiguiente

$$d_H^{(\mathcal{Z}, \partial)}(f_{\bar{\xi}}(\mathcal{V}), f_{\bar{\xi}^k}(\mathcal{V})) = d_H^{(\mathcal{Z}, \partial)}(\{\xi_1, \dots, \xi_n\}, \{\xi_1^k, \dots, \xi_n^k\}).$$

Sea $\epsilon > 0$ dado. Entonces, como $\rho(\xi_i^k, \xi_i) \rightarrow 0$, $k \rightarrow \infty$ para todo $i = 1, \dots, n$, existe $k(\epsilon)$ tal que si $k \geq k(\epsilon)$, tenemos que $\rho(\xi_i^k, \xi_i) < \epsilon$. Por lo tanto la distancia de Hausdorff con respecto a δ que es ρ entre $\{\xi_1, \dots, \xi_n\}$ y $\{\xi_1^k, \dots, \xi_n^k\}$ es menor o igual que ϵ si $k \geq k(\epsilon)$. En otras palabras

$$d_{GH}(\mathcal{V}_{\bar{\xi}}, \mathcal{V}_{\bar{\xi}^k}) \leq \epsilon,$$

para todo $k \geq k(\epsilon)$. □

6.4. Algunos ejemplos

EJEMPLO 6.4.1. Volvemos a considerar las situaciones descritas en los Ejemplos 4.2.1 y 5.4.1. En el Ejemplo 4.2.1 encontramos que la distancia de Hausdorff entre el conjunto \mathcal{V}_1 y \mathcal{V}_2 es aproximadamente 1.58. En el ejemplo 5.4.1 obtuvimos que la distancia de Gromov-Lipschitz entre (\mathcal{V}_1, d_1) y (\mathcal{V}_2, d_2) es aproximadamente 0.69. Es claro que si consideramos como (Y, δ) el espacio \mathbb{R}^2 con su métrica usual y f_1, f_2 las inclusiones de \mathcal{V}_1 y \mathcal{V}_2 en \mathbb{R}^2 , resulta también que la distancia de Gromov-Hausdorff entre (\mathcal{V}_1, d_1) y (\mathcal{V}_2, d_2) es menor o igual a 1.58. Pero resulta claro que esta elección de f_1 y f_2 está lejos de ser

óptima. Podemos obtener mejores estimaciones para d_{GH} , en efecto si f_1 es nuevamente la inclusión de \mathcal{V}_1 en \mathbb{R}^2 y si f_2 es la traslación de \mathcal{V}_2 dada por el vector $(-\frac{1}{2}, 0)$, obtenemos $\mathcal{V}_2^* = \{(\frac{3}{2}, 1), (\frac{5}{2}, 2), (1, \frac{3}{2}), (\frac{1}{2}, 3)\}$ y ahora la distancia de Hausdorff entre $f_1(\mathcal{V}_1)$ y $f_2(\mathcal{V}_2) = \mathcal{V}_2^*$ es del orden de 1.12. No obstante como también las rotaciones son isometrías en el plano, podemos encontrar mejores estimaciones para $d_{GH}((\mathcal{V}_1, d_1), (\mathcal{V}_2, d_2))$. Si tomamos ahora f_1 la traslación en $(0, -1)$, entonces $f_1(\mathcal{V}_1) = \mathcal{V}_1^* = \{(0, 0), (0, 1), (2, 2), (1, 0)\}$. Para construir una f_2 trasladamos primero \mathcal{V}_2 en $(-2, -1)$ y luego tomamos una rotación en un ángulo de $-\frac{3\pi}{8}$ lo cual nos da la matriz $\rho = \begin{pmatrix} \frac{\sqrt{2-\sqrt{2}}}{2} & \frac{\sqrt{2+\sqrt{2}}}{2} \\ -\frac{\sqrt{2+\sqrt{2}}}{2} & \frac{\sqrt{2-\sqrt{2}}}{2} \end{pmatrix}$. Así tomamos $f_2 = \rho \circ \tau$, donde τ es la traslación. En otras palabras

$$\begin{aligned} f_2(x, y) &= \begin{pmatrix} \frac{\sqrt{2-\sqrt{2}}}{2} & \frac{\sqrt{2+\sqrt{2}}}{2} \\ -\frac{\sqrt{2+\sqrt{2}}}{2} & \frac{\sqrt{2-\sqrt{2}}}{2} \end{pmatrix} \cdot \begin{pmatrix} x-2 \\ y-1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{\sqrt{2-\sqrt{2}}}{2} \cdot (x-2) + \frac{\sqrt{2+\sqrt{2}}}{2} \cdot (y-1) \\ \frac{\sqrt{2-\sqrt{2}}}{2} \cdot (-x+2) + \frac{\sqrt{2+\sqrt{2}}}{2} \cdot (y-1) \end{pmatrix}. \end{aligned}$$

Por consiguiente $f(\mathcal{V}_2) = \mathcal{V}_2^* = \{(0, 0), (\frac{\sqrt{2-\sqrt{2}}+\sqrt{2+\sqrt{2}}}{2}, \frac{-\sqrt{2+\sqrt{2}}+\sqrt{2-\sqrt{2}}}{2}), (\frac{-\sqrt{2-\sqrt{2}}+\sqrt{2+\sqrt{2}}}{4}, \frac{\sqrt{2-\sqrt{2}}+\sqrt{2+\sqrt{2}}}{4}), (-\frac{\sqrt{2-\sqrt{2}}}{2} + \sqrt{2+\sqrt{2}}, \frac{\sqrt{2+\sqrt{2}}}{2} + \sqrt{2-\sqrt{2}})\}$ como se muestra en la Figura 6.4.4.

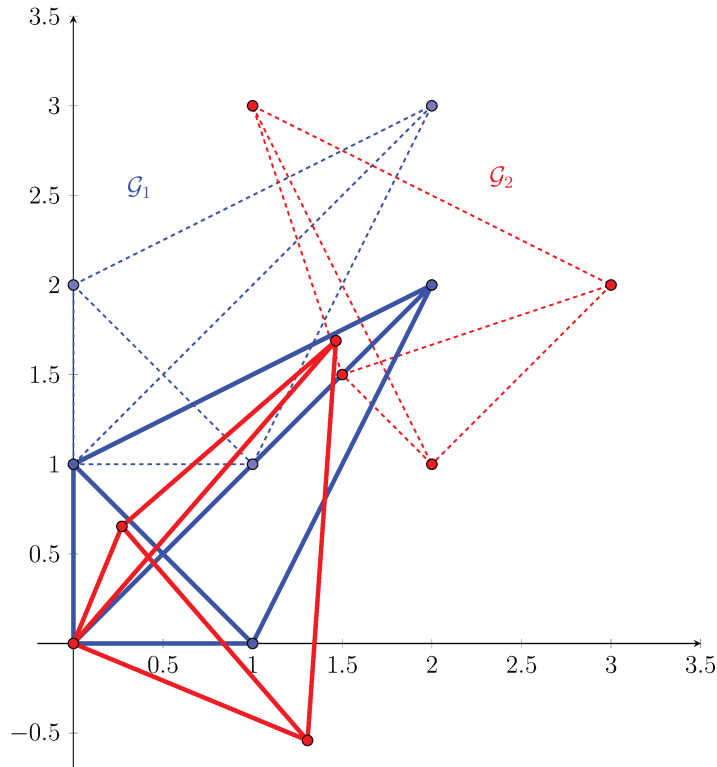


FIGURA 6.4.4. \mathcal{V}_1^* y \mathcal{V}_2^* inmersiones isométricas más cercanas al óptimo.

Aplicando el algoritmo del Capítulo 4 en la Sección 4.2, la distancia de Hausdorff entre el conjunto \mathcal{V}_1^* y \mathcal{V}_2^* es aproximadamente 0.62. Por lo tanto $d_{GH}((\mathcal{V}_1, d_1), (\mathcal{V}_2, d_2)) \leq 0.62$.

EJEMPLO 6.4.2. En este ejemplo vamos a considerar dos grafos con cardinales distintos. Sean $\mathcal{V}_1 = \{(0, 0), (0, 4), (4, 4), (4, 0)\}$ y $\mathcal{V}_2 = \{(0, 0), (0, 4), (4, 0)\}$ dos conjuntos en \mathbb{R}^2 con $\#(\mathcal{V}_1) \neq \#(\mathcal{V}_2)$. Con los conceptos introducidos en la Proposición 6.2.1 la distancia de Gromov-Hausdorff es menor o igual que la distancia de Hausdorff entre los conjuntos \mathcal{V}_1 y \mathcal{V}_2 que es 4. Ahora si consideramos (Y, δ) el espacio en \mathbb{R}^2 con su métrica usual y f_1 la inclusión en \mathcal{V}_1 , f_2 una rotación y traslación como nos muestra la Figura 6.4.5 resultan isométricos los conjuntos \mathcal{V}_2 con $f_2(\mathcal{V}_2) = \mathcal{V}_2^* = \{(2(1 - \sqrt{2}), 2), (2(1 + \sqrt{2}), 2), (2, 2(1 - \sqrt{2}))\}$ por lo tanto la distancia de Gromov-Hausdorff entre el grafo \mathcal{G}_1 y el grafo \mathcal{G}_2 es menor o igual que la distancia de Hausdorff entre \mathcal{V}_1 y \mathcal{V}_2^* que es aproximadamente 2.17.

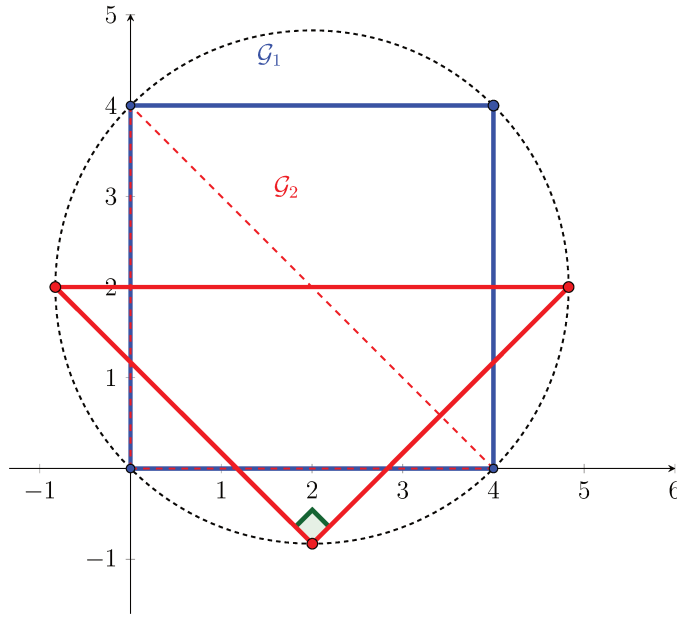


FIGURA 6.4.5. Distancia de Gromov-Hausdorff entre los grafos con diferentes \mathcal{G}_1 y \mathcal{G}_2 números de vértices.

EJEMPLO 6.4.3. Sean (\mathcal{V}_1, d_1) y (\mathcal{V}_2, d_2) dos grafos métricos finitos. Sea $\mathcal{V}_1 \sqcup \mathcal{V}_2$ la unión disjunta de los espacios \mathcal{V}_1 y \mathcal{V}_2 definida como $Y = \mathcal{V}_1 \sqcup \mathcal{V}_2 = (\mathcal{V}_1 \times \{1\}) \cup (\mathcal{V}_2 \times \{2\}) := \tilde{\mathcal{V}}_1 \cup \tilde{\mathcal{V}}_2$. Entonces

$$\tilde{v}_1 \in \tilde{\mathcal{V}}_1 = \mathcal{V}_1 \times \{1\}, \tilde{v}_1 = (v_1, 1) \text{ con } v_1 \in \mathcal{V}_1;$$

$$\tilde{v}_2 \in \tilde{\mathcal{V}}_2 = \mathcal{V}_2 \times \{2\}, \tilde{v}_2 = (v_2, 2) \text{ con } v_2 \in \mathcal{V}_2.$$

Sean $\tilde{x}, \tilde{y} \in Y$, entonces

$$\delta(\tilde{x}, \tilde{y}) = \begin{cases} d_1(\tilde{x}, \tilde{y}) & \text{si } \tilde{x} \in \tilde{\mathcal{V}}_1, \tilde{y} \in \tilde{\mathcal{V}}_1; \\ d_2(\tilde{x}, \tilde{y}) & \text{si } \tilde{x} \in \tilde{\mathcal{V}}_2, \tilde{y} \in \tilde{\mathcal{V}}_2; \\ \text{máx}\{\text{diam}\mathcal{V}_1, \text{diam}\mathcal{V}_2\} & \text{si } \tilde{x} \in \tilde{\mathcal{V}}_1, \tilde{y} \in \tilde{\mathcal{V}}_2 \text{ o } \tilde{x} \in \tilde{\mathcal{V}}_2, \tilde{y} \in \tilde{\mathcal{V}}_1, \end{cases}$$

es una métrica. Hay que probar $M_1)$, $M_2)$, $M_3)$ y $M_4)$ del Capítulo 2 de la Sección 2.5. Las propiedades $M_1)$, $M_2)$ y $M_3)$ son obvias. Para la desigualdad triangular, $M_4)$, consideramos \tilde{x} , \tilde{y} y $\tilde{z} \in Y$. Si \tilde{x} , \tilde{y} y \tilde{z} están en $\tilde{\mathcal{V}}_1$ o \tilde{x} , \tilde{y} y \tilde{z} están en $\tilde{\mathcal{V}}_2$ la desigualdad triangular para δ se sigue de la desigualdad triangular para la correspondiente d_i , $i = 1, 2$. En el caso que $\tilde{x}, \tilde{y} \in \tilde{\mathcal{V}}_1$ y $\tilde{z} \in \tilde{\mathcal{V}}_2$

$$\begin{aligned} \delta(\tilde{x}, \tilde{z}) &= \text{máx}\{\text{diam}\mathcal{V}_1, \text{diam}\mathcal{V}_2\} \\ &\leq \text{máx}\{\text{diam}\mathcal{V}_1, \text{diam}\mathcal{V}_2\} + d_1(\tilde{x}, \tilde{y}) \\ &= d_1(\tilde{x}, \tilde{y}) + \delta(\tilde{y}, \tilde{z}) \\ &= \delta(\tilde{x}, \tilde{y}) + \delta(\tilde{y}, \tilde{z}). \end{aligned}$$

Análogamente para los casos restantes. Por consiguiente, por la Proposición 6.2.1, tendremos que $d_{GH}((\mathcal{V}_1, d_1), (\mathcal{V}_2, d_2)) \leq \delta_H(\tilde{\mathcal{V}}_1, \tilde{\mathcal{V}}_2)$ donde δ_H es la distancia de Hausdorff en el espacio métrico (Y, δ) definido arriba. Con este esquema estimamos nuevamente la distancia de Gromov-Hausdorff del ejemplo contenido en 6.4.1. Primero obtenemos la unión disjunta de los conjuntos \mathcal{V}_1 y \mathcal{V}_2 como se describe arriba. Ahora calculamos los diámetros de cada conjunto para obtener el máximo, entonces el $\text{diam}\mathcal{V}_1 = 2.83$ y el $\text{diam}\mathcal{V}_2 = 2.24$ podemos ver que el máximo es $\text{diam}\mathcal{V}_1$, por ultimo estamos en condiciones de calcular la distancia de Hausdorff en el espacio métrico (Y, δ) . La distancia de Hausdorff de $\tilde{\mathcal{V}}_1$ y $\tilde{\mathcal{V}}_2$ con la métrica δ es $\delta_H(\tilde{\mathcal{V}}_1, \tilde{\mathcal{V}}_2) = 2.83$. Es claro que no es una buena cota para la distancia de Gromov-Hausdorff ya que en el Ejemplo 6.4.1 obtuvimos una estimación mucho mejor para: $d_{GH}((\mathcal{V}_1, d_1), (\mathcal{V}_2, d_2)) \leq 0.62$.

6.5. Cálculo de distancia de Gromov-Hausdorff usando correspondencias

En contextos discretos y computacionalmente accesibles, resulta conveniente una reformulación de d_{GH} debida a [Mí2].

TEOREMA 6.5.1. *Dados dos espacios métricos (X_i, d_i) , $i = 1, 2$, se tiene que*

$$d_{GH}((X_1, d_1), (X_2, d_2)) = \frac{1}{2} \inf\{\text{dist}(R) : R \text{ correspondencia entre } X_1 \text{ y } X_2\},$$

donde

a) *R es una correspondencia entre X_1 y X_2 si $R \subset X_1 \times X_2$ y satisface que para todo $x_1 \in X_1$, existe $x_2 \in X_2$ tal que $(x_1, x_2) \in R$ y también para todo $x_2 \in X_2$, existe $x_1 \in X_1$ tal que $(x_1, x_2) \in R$;*

b) $\text{dist}(R) = \sup_{\substack{(x_1, x_2) \in R \\ (x'_1, x'_2) \in R}} |d_1(x_1, x'_1) - d_2(x_2, x'_2)|$ es la distorsión de la correspondencia R .

En algunos casos de espacios de bajo número de elementos es posible diseñar a partir del teorema anterior un algoritmo de cálculo de d_{GH} , que es el siguiente. En primer lugar se considera el conjunto de partes $\mathcal{P}(X_1 \times X_2)$. Estamos suponiendo que tanto X_1 como X_2 son finitos. Sean $N_1 = \#(X_1)$ y $N_2 = \#(X_2)$. Entonces $\#(X_1 \times X_2) = N_1 N_2$ y $\#(\mathcal{P}(X_1 \times X_2)) = 2^{N_1 N_2}$. Ahora consideramos los dos proyectores $\Pi_1 : X_1 \times X_2 \rightarrow X_1$, $\Pi_1(x_1, x_2) = x_1$ y $\Pi_2 : X_1 \times X_2 \rightarrow X_2$, $\Pi_2(x_1, x_2) = x_2$.

LEMA 6.5.2. *$R \in \mathcal{P}(X_1 \times X_2)$ es correspondencia si y sólo si $\#(\Pi_1(R)) = N_1$ y $\#(\Pi_2(R)) = N_2$.*

DEMOSTRACIÓN. Es claro que R es correspondencia si y sólo si $\Pi_1(R) = X_1$ y $\Pi_2(R) = X_2$. Pero como estamos en contextos finitos $\Pi_i(R) = X_i$ para $i = 1, 2$ si y sólo si $\#(\Pi_i(R)) = \#(X_i) = N_i$ □

El algoritmo usa el lema para descartar de $\mathcal{P}(X_1 \times X_2)$ todas las relaciones que no son correspondencias. A las restantes calculamos la distorsión

$$\text{dist}(R) = \sup_{\substack{(x_1, x_2) \in R \\ (x'_1, x'_2) \in R}} |d_1(x_1, x'_1) - d_2(x_2, x'_2)|.$$

Finalmente se calcula la mitad del mínimo de todas estas distorsiones para las R que son correspondencias.

Aplicamos este algoritmo a dos ejemplos anteriores, el Ejemplo 6.4.1 y Ejemplo 6.4.2. Implementando en Python el algoritmo anterior provee ahora

$$d_{GH}((\mathcal{V}_{6.4.1}^1, d_{6.4.1}^1), (\mathcal{V}_{6.4.1}^2, d_{6.4.1}^2)) \simeq 0.327$$

y

$$d_{GH}((\mathcal{V}_{6.4.2}^1, d_{6.4.2}^1), (\mathcal{V}_{6.4.2}^2, d_{6.4.2}^2)) = 2.$$

En ambos casos el valor para d_{GH} es estrictamente menor que las mejores cotas que habíamos encontrado con argumentos geométricos, es decir 0.62 y 2.17 respectivamente.

6.6. Aplicación de la distancia de Gromov-Hausdorff al transporte público en AMBA

En el contexto de los datos SUBE de transporte en AMBA consideramos la situación descrita en el Ejemplo 4.2.2 en la que \mathcal{V} es el conjunto de los 41 distritos de AMBA y la métrica es la distancia difusiva correspondiente a $t = 100$. Aquí tomaremos el grafo métrico $\mathcal{G}_1 = (\mathcal{V}_1, d_1)$ igual al grafo de AMBA. Nos proponemos ver cuán representativo de \mathcal{G}_1 es el grafo \mathcal{G}_2 que se obtiene usando los 16 agrupamientos por proximidad geográfica del Ejemplo 4.2.2, de la manera que a continuación describimos. Recordamos que agrupábamos los distritos de AMBA, $A = \bigcup_{j=1}^{16} A_j$ de la siguiente manera

$$A_1 = \{28, 3, 12\} = \{\text{Zárate, Campana, San Fernando}\};$$

$$A_2 = \{11, 4, 27\} = \{\text{Pilar, Escobar, Tigre}\};$$

$$A_3 = \{32, 9\} = \{\text{Exaltación de la Cruz, Luján}\};$$

$$A_4 = \{33, 6, 23\} = \{\text{General Las Heras, General Rodríguez, Marcos Paz}\};$$

$$A_5 = \{17, 13, 16\} = \{\text{Cañuelas, San Vicente, Brandsen}\};$$

$$A_6 = \{22, 31, 2\} = \{\text{La Plata, Ensenada, Berisso}\};$$

$$A_7 = \{29, 19, 25, 1\} = \{\text{Berazategui, Florencio Varela, Presidente Perón, Almirante Brown}\};$$

$$A_8 = \{8, 15, 39\} = \{\text{Lanús, Avellaneda, Quilmes}\};$$

$$A_9 = \{36, 18, 5\} = \{\text{Lomas de Zamora, Esteban Echeverría, Ezeiza}\};$$

$$A_{10} = \{35\} = \{\text{La Matanza}\};$$

$$A_{11} = \{30\} = \{\text{CABA}\};$$

- $A_{12} = \{20, 41\} = \{\text{General San Martín, Tres de Febrero}\};$
- $A_{13} = \{7, 34, 38\} = \{\text{Ituzzaingó, Hurlingham, Morón}\};$
- $A_{14} = \{24, 10\} = \{\text{Merlo, Moreno}\};$
- $A_{15} = \{21, 37, 40\} = \{\text{José C. Paz, Malvinas Argentinas, San Miguel}\};$
- $A_{16} = \{26, 14\} = \{\text{San Isidro, Vicente López}\}.$

Ahora de cada uno de ellos elegimos el distrito de mayor población (censo 2022) lo cual nos da el conjunto $\mathcal{V}_2 = \{12, 27, 9, 6, 13, 22, 1, 39, 36, 35, 30, 20, 38, 24, 37, 26\}$. Como $\mathcal{V}_2 \subset \mathcal{V}_1$, tenemos una isometría trivial dada por la inclusión de \mathcal{V}_2 en \mathcal{V}_1 si a \mathcal{V}_2 lo equipamos con la métrica que hereda, que es la métrica difusiva (generada por el Laplaciano en \mathcal{V}_1) para tiempo $t = 100$. Así, $(Y, \delta) = (\mathcal{V}_1, d_1)$ y por lo tanto $d_{GH}(\mathcal{G}_1, \mathcal{G}_2) \leq \delta_H(\mathcal{V}_1, \mathcal{V}_2)$. Usando otra vez el algoritmo del Capítulo 4, tenemos que $d_{GH}(\mathcal{G}_1, \mathcal{G}_2) = 0.993$. Esta distancia es grande en relación con los valores que muestra la distancia difusiva en tiempo $t = 100$ en la Figura 6.6.6.

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 0.37 | 0.84 | 0.92 | 0 | 0.54 | 0.7 | 0.52 | 0.36 | 0.93 | 0.34 | 0.31 | 0.44 | 0.79 | 0.32 | 0.37 | 1.04 | 0.95 | 0.4 | 0.39 | 0.35 | 0.34 | 0.5 | 0.89 | 0.36 | 0.64 | 0.32 | 0.29 | 0.99 | 0.43 | 0.34 | 0.9 | 1.02 | 1.05 | 0.5 | 0.35 | 0.36 | 0.32 | 0.36 | 0.37 | 0.34 | 0.35 | |
| 0.38 | 0.86 | 0.97 | 0.54 | 0 | 0.74 | 0.55 | 0.38 | 0.96 | 0.41 | 0.48 | 0.53 | 0.8 | 0.4 | 0.39 | 1.06 | 0.93 | 0.32 | 0.41 | 0.41 | 0.44 | 0.53 | 0.91 | 0.41 | 0.65 | 0.42 | 0.43 | 1.02 | 0.45 | 0.38 | 0.92 | 1.04 | 1.07 | 0.55 | 0.39 | 0.37 | 0.45 | 0.41 | 0.4 | 0.43 | 0.4 | |
| 0.63 | 0.98 | 1.07 | 0.7 | 0.74 | 0 | 0.71 | 0.63 | 1.02 | 0.58 | 0.63 | 0.7 | 0.94 | 0.62 | 0.63 | 1.15 | 1.07 | 0.65 | 0.64 | 0.63 | 0.63 | 0.71 | 1.02 | 0.61 | 0.81 | 0.63 | 0.63 | 1.11 | 0.67 | 0.62 | 1.03 | 1.14 | 1.16 | 0.71 | 0.62 | 0.63 | 0.64 | 0.62 | 0.63 | 0.62 | 0.62 | |
| 0.39 | 0.85 | 0.96 | 0.52 | 0.55 | 0.71 | 0 | 0.38 | 0.95 | 0.35 | 0.45 | 0.51 | 0.8 | 0.38 | 0.39 | 1.05 | 0.96 | 0.41 | 0.41 | 0.38 | 0.41 | 0.52 | 0.69 | 0.33 | 0.65 | 0.4 | 0.41 | 1.01 | 0.45 | 0.37 | 0.92 | 1.04 | 1.06 | 0.49 | 0.36 | 0.38 | 0.42 | 0.32 | 0.4 | 0.39 | 0.37 | |
| 0.04 | 0.77 | 0.89 | 0.36 | 0.38 | 0.63 | 0.38 | 0 | 0.88 | 0.13 | 0.27 | 0.35 | 0.7 | 0.11 | 0.04 | 0.99 | 0.88 | 0.14 | 0.13 | 0.12 | 0.2 | 0.35 | 0.83 | 0.12 | 0.51 | 0.15 | 0.18 | 0.94 | 0.21 | 0.06 | 0.84 | 0.97 | 1 | 0.38 | 0.07 | 0.03 | 0.21 | 0.12 | 0.07 | 0.17 | 0.1 | |
| 0.88 | 1.15 | 1.23 | 0.93 | 0.96 | 1.02 | 0.95 | 0.88 | 0 | 0.87 | 0.9 | 0.93 | 1.11 | 0.88 | 0.88 | 1.3 | 1.23 | 0.89 | 0.89 | 0.88 | 0.89 | 0.93 | 1.19 | 0.88 | 1.01 | 0.88 | 0.88 | 1.26 | 0.9 | 0.87 | 1.19 | 1.28 | 1.31 | 0.94 | 0.87 | 0.88 | 0.89 | 0.88 | 0.88 | 0.88 | 0.88 | 0.88 |
| 0.14 | 0.77 | 0.89 | 0.34 | 0.41 | 0.58 | 0.35 | 0.13 | 0.87 | 0 | 0.23 | 0.33 | 0.72 | 0.09 | 0.14 | 0.99 | 0.89 | 0.2 | 0.19 | 0.11 | 0.14 | 0.37 | 0.62 | 0.09 | 0.54 | 0.13 | 0.14 | 0.94 | 0.26 | 0.08 | 0.84 | 0.97 | 1 | 0.36 | 0.08 | 0.12 | 0.17 | 0.09 | 0.15 | 0.11 | 0.09 | |
| 0.28 | 0.81 | 0.91 | 0.31 | 0.48 | 0.63 | 0.45 | 0.27 | 0.9 | 0.23 | 0 | 0.39 | 0.75 | 0.23 | 0.28 | 1.01 | 0.92 | 0.31 | 0.31 | 0.26 | 0.24 | 0.44 | 0.86 | 0.27 | 0.59 | 0.24 | 0.23 | 0.97 | 0.35 | 0.25 | 0.87 | 0.99 | 1.03 | 0.44 | 0.25 | 0.27 | 0.23 | 0.26 | 0.29 | 0.24 | 0.25 | |
| 0.35 | 0.84 | 0.94 | 0.44 | 0.53 | 0.7 | 0.51 | 0.35 | 0.93 | 0.33 | 0.39 | 0 | 0.79 | 0.3 | 0.35 | 1.04 | 0.94 | 0.38 | 0.38 | 0.33 | 0.35 | 0.49 | 0.89 | 0.35 | 0.63 | 0.27 | 0.25 | 0.99 | 0.42 | 0.33 | 0.9 | 1.02 | 1.05 | 0.5 | 0.33 | 0.35 | 0.34 | 0.35 | 0.36 | 0.34 | 0.34 | |
| 0.69 | 1.03 | 1.12 | 0.79 | 0.8 | 0.94 | 0.8 | 0.7 | 1.11 | 0.72 | 0.75 | 0.79 | 0 | 0.71 | 0.71 | 1.18 | 1.12 | 0.72 | 0.72 | 0.72 | 0.73 | 0.78 | 1.08 | 0.72 | 0.84 | 0.72 | 0.73 | 1.16 | 0.74 | 0.7 | 1.09 | 1.19 | 1.21 | 0.8 | 0.71 | 0.7 | 0.74 | 0.72 | 0.71 | 0.73 | 0.71 | |
| 0.11 | 0.77 | 0.89 | 0.32 | 0.4 | 0.62 | 0.38 | 0.1 | 0.88 | 0.09 | 0.23 | 0.3 | 0.71 | 0 | 0.11 | 0.99 | 0.89 | 0.18 | 0.17 | 0.08 | 0.15 | 0.36 | 0.82 | 0.11 | 0.53 | 0.07 | 0.1 | 0.94 | 0.24 | 0.06 | 0.84 | 0.97 | 1 | 0.37 | 0.07 | 0.09 | 0.14 | 0.11 | 0.13 | 0.12 | 0.08 | |
| 0.06 | 0.77 | 0.89 | 0.37 | 0.39 | 0.63 | 0.39 | 0.04 | 0.88 | 0.14 | 0.28 | 0.35 | 0.71 | 0.11 | 0 | 0.99 | 0.89 | 0.15 | 0.12 | 0.13 | 0.21 | 0.35 | 0.83 | 0.13 | 0.52 | 0.16 | 0.19 | 0.94 | 0.2 | 0.07 | 0.84 | 0.98 | 1 | 0.38 | 0.08 | 0.06 | 0.21 | 0.13 | 0.06 | 0.18 | 0.11 | |
| 0.99 | 1.23 | 1.31 | 1.04 | 1.06 | 1.15 | 1.05 | 0.99 | 1.3 | 0.99 | 1.01 | 1.04 | 1.18 | 0.99 | 0.99 | 0 | 1.31 | 1 | 1 | 0.99 | 1 | 1.03 | 1.27 | 0.99 | 1.11 | 0.99 | 1 | 1.34 | 1.01 | 0.99 | 1.27 | 1.36 | 1.38 | 1.05 | 0.99 | 0.99 | 1 | 0.99 | 0.99 | 0.99 | 0.99 | |
| 0.88 | 1.15 | 1.24 | 0.95 | 0.99 | 1.07 | 0.96 | 0.88 | 1.23 | 0.89 | 0.92 | 0.94 | 1.12 | 0.89 | 0.89 | 1.31 | 0 | 0.89 | 0.9 | 0.89 | 0.9 | 0.94 | 1.19 | 0.89 | 1.02 | 0.89 | 0.9 | 1.27 | 0.91 | 0.88 | 1.2 | 1.29 | 1.32 | 0.95 | 0.88 | 0.88 | 0.9 | 0.89 | 0.89 | 0.89 | 0.89 | |
| 0.14 | 0.78 | 0.9 | 0.4 | 0.32 | 0.65 | 0.41 | 0.14 | 0.89 | 0.2 | 0.31 | 0.38 | 0.72 | 0.18 | 0.15 | 1 | 0.89 | 0 | 0.19 | 0.25 | 0.39 | 0.84 | 0.19 | 0.54 | 0.21 | 0.24 | 0.95 | 0.27 | 0.15 | 0.85 | 0.98 | 1.01 | 0.41 | 0.15 | 0.12 | 0.26 | 0.19 | 0.17 | 0.23 | 0.18 | | |
| 0.13 | 0.77 | 0.9 | 0.39 | 0.41 | 0.64 | 0.41 | 0.13 | 0.89 | 0.19 | 0.31 | 0.38 | 0.72 | 0.17 | 0.12 | 1 | 0.9 | 0.2 | 0 | 0.18 | 0.25 | 0.34 | 0.84 | 0.18 | 0.54 | 0.2 | 0.23 | 0.95 | 0.17 | 0.13 | 0.84 | 0.98 | 1.01 | 0.4 | 0.15 | 0.13 | 0.25 | 0.18 | 0.08 | 0.22 | 0.17 | |
| 0.13 | 0.78 | 0.89 | 0.35 | 0.41 | 0.63 | 0.38 | 0.12 | 0.88 | 0.11 | 0.26 | 0.33 | 0.72 | 0.08 | 0.13 | 0.99 | 0.89 | 0.19 | 0.18 | 0 | 0.18 | 0.37 | 0.83 | 0.12 | 0.54 | 0.11 | 0.15 | 0.95 | 0.25 | 0.09 | 0.85 | 0.98 | 1 | 0.36 | 0.09 | 0.12 | 0.18 | 0.11 | 0.15 | 0.14 | 0.06 | |
| 0.21 | 0.79 | 0.9 | 0.34 | 0.44 | 0.63 | 0.41 | 0.2 | 0.89 | 0.14 | 0.24 | 0.35 | 0.73 | 0.15 | 0.21 | 1 | 0.9 | 0.25 | 0.25 | 0.18 | 0 | 0.4 | 0.84 | 0.19 | 0.56 | 0.17 | 0.17 | 0.95 | 0.3 | 0.17 | 0.86 | 0.98 | 1.01 | 0.39 | 0.17 | 0.2 | 0.16 | 0.19 | 0.22 | 0.1 | 0.17 | |
| 0.35 | 0.63 | 0.95 | 0.5 | 0.53 | 0.71 | 0.52 | 0.35 | 0.93 | 0.37 | 0.44 | 0.49 | 0.78 | 0.36 | 0.35 | 1.03 | 0.94 | 0.39 | 0.34 | 0.37 | 0.4 | 0 | 0.89 | 0.37 | 0.63 | 0.38 | 0.39 | 0.99 | 0.36 | 0.35 | 0.77 | 1.02 | 1.05 | 0.51 | 0.35 | 0.35 | 0.41 | 0.37 | 0.33 | 0.39 | 0.36 | |
| 0.83 | 1.11 | 1.2 | 0.89 | 0.91 | 1.02 | 0.89 | 0.83 | 1.19 | 0.82 | 0.86 | 0.89 | 1.08 | 0.82 | 0.83 | 1.27 | 1.19 | 0.84 | 0.84 | 0.83 | 0.84 | 0.89 | 0 | 0.8 | 0.97 | 0.83 | 0.84 | 1.23 | 0.85 | 0.82 | 1.16 | 1.26 | 1.27 | 0.9 | 0.82 | 0.83 | 0.84 | 0.82 | 0.83 | 0.83 | 0.82 | |
| 0.13 | 0.77 | 0.89 | 0.36 | 0.41 | 0.61 | 0.39 | 0.12 | 0.88 | 0.09 | 0.27 | 0.35 | 0.72 | 0.11 | 0.13 | 0.99 | 0.89 | 0.19 | 0.18 | 0.12 | 0.19 | 0.37 | 0.8 | 0 | 0.54 | 0.15 | 0.18 | 0.95 | 0.25 | 0.09 | 0.85 | 0.98 | 1 | 0.36 | 0.07 | 0.12 | 0.21 | 0.05 | 0.15 | 0.16 | 0.11 | |
| 0.48 | 0.92 | 1.03 | 0.64 | 0.65 | 0.81 | 0.65 | 0.51 | 1.01 | 0.54 | 0.59 | 0.63 | 0.84 | 0.53 | 0.52 | 1.11 | 1.02 | 0.54 | 0.54 | 0.54 | 0.56 | 0.63 | 0.97 | 0.54 | 0 | 0.54 | 0.55 | 1.07 | 0.57 | 0.52 | 0.98 | 1.1 | 1.12 | 0.65 | 0.52 | 0.53 | 0.57 | 0.54 | 0.52 | 0.55 | 0.53 | |
| 0.16 | 0.78 | 0.89 | 0.32 | 0.42 | 0.63 | 0.4 | 0.15 | 0.88 | 0.13 | 0.24 | 0.27 | 0.72 | 0.07 | 0.16 | 0.99 | 0.89 | 0.21 | 0.2 | 0.11 | 0.17 | 0.38 | 0.83 | 0.15 | 0.54 | 0 | 0.09 | 0.95 | 0.27 | 0.11 | 0.85 | 0.98 | 1.01 | 0.38 | 0.12 | 0.14 | 0.16 | 0.15 | 0.17 | 0.15 | 0.12 | |
| 0.18 | 0.78 | 0.89 | 0.29 | 0.43 | 0.63 | 0.41 | 0.18 | 0.88 | 0.14 | 0.23 | 0.25 | 0.73 | 0.1 | 0.19 | 1 | 0.9 | 0.24 | 0.23 | 0.15 | 0 | 0.89 | 0.37 | 0.63 | 0.38 | 0.39 | 0.99 | 0.36 | 0.35 | 0.77 | 1.02 | 1.05 | 0.51 | 0.35 | 0.35 | 0.41 | 0.37 | 0.33 | 0.39 | 0.36 | | |
| 0.94 | 1.2 | 1.21 | 0.99 | 1.02 | 1.11 | 1.01 | 0.94 | 1.26 | 0.94 | 0.97 | 0.99 | 1.16 | 0.94 | 0.94 | 1.34 | 1.27 | 0.95 | 0.95 | 0.95 | 0.99 | 1.23 | 0.95 | 1.07 | 0.95 | 0.95 | 0 | 0.97 | 0.94 | 1.24 | 1.33 | 1.35 | 1.01 | 0.94 | 0.94 | 0.96 | 0.94 | 0.95 | 0.95 | 0.94 | | |
| 0.21 | 0.78 | 0.92 | 0.43 | 0.45 | 0.67 | 0.45 | 0.21 | 0.9 | 0.26 | 0.35 | 0.42 | 0.74 | 0.24 | 0.2 | 1.01 | 0.91 | 0.27 | 0.17 | 0.25 | 0.3 | 0.36 | 0.85 | 0.25 | 0.57 | 0.27 | 0.29 | 0.97 | 0 | 0.21 | 0.86 | 1 | 1.02 | 0.44 | 0.22 | 0.21 | 0.31 | 0.25 | 0.15 | 0.28 | 0.24 | |
| 0.07 | 0.65 | 0.89 | 0.34 | 0.38 | 0.62 | 0.37 | 0.06 | 0.87 | 0.08 | 0.25 | 0.33 | 0.7 | 0.06 | 0.07 | 0.99 | 0.88 | 0.15 | 0.13 | 0.09 | 0.17 | 0.35 | 0.82 | 0.09 | 0.52 | 0.11 | 0.14 | 0.94 | 0.21 | 1 | 0.84 | 0.97 | 1 | 0.36 | 0.03 | 0.05 | 0.18 | 0.08 | 0.09 | 0.13 | 0.07 | |
| 0.84 | 1.07 | 1.2 | 0.92 | 0.92 | 1.03 | 0.92 | 0.84 | 1.19 | 0.84 | 0.87 | 0.9 | 1.09 | 0.84 | 0.84 | 1.27 | 1.2 | 0.85 | 0.84 | 0.85 | 0.86 | 1.16 | 1.16 | 0.85 | 0.98 | 0.85 | 0.85 | 1.24 | 0.86 | 0.84 | 0 | 1.26 | 1.29 | 0.91 | 0.84 | 0.84 | 0.86 | 0.84 | 0.84 | 0.85 | 0.84 | |
| 0.97 | 1.22 | 1.29 | 1.02 | 1.04 | 1.14 | 1.04 | 0.97 | 1.28 | 0.97 | 0.99 | 1.02 | 1.19 | 0.97 | 0.98 | 1.36 | 1.29 | 0.98 | 0.98 | 0.98 | 1.02 | 1.26 | 0.98 | 1.1 | 0.98 | 0.98 | 1.33 | 1 | 0.97 | 1.26 | 0 | 1.37 | 1.03 | 0.97 | 0.97 | 0.98 | 0.97 | 0.98 | 0.98 | 0.97 | | |
| 1 | 1.24 | 1.32 | 1.05 | 1.07 | 1.16 | 1.06 | 1 | 1.31 | 1 | 1.03 | 1.05 | 1.21 | 1 | 1 | 1.38 | 1.32 | 1.01 | 1.01 | 1 | 1.01 | 1.05 | 1.27 | 1 | 1.12 | 1.01 | 1.01 | 1.35 | 1.02 | 1 | 1.29 | 1.37 | 0 | 1.06 | 1 | 1 | 1.01 | 1 | 1 | 1.01 | 1 | |
| 0.58 | 0.85 | 0.96 | 0.5 | 0.55 | 0.71 | 0.49 | 0.38 | 0.94 | 0.36 | 0.44 | 0.5 | 0.8 | 0.37 | 0.38 | 1.04 | 0.95 | 0.41 | 0.4 | 0.36 | 0.39 | 0.51 | 0.9 | 0.36 | 0.65 | 0.38 | 0.39 | 1.01 | 0.44 | 0.36 | 0.91 | 1.03 | 1.06 | 0 | 0.35 | 0.37 | 0.4 | 0.33 | 0.39 | 0.36 | 0.35 | |
| 0.08 | 0.77 | 0.89 | 0.35 | 0.39 | 0.62 | 0.36 | 0.07 | 0.87 | 0.08 | 0.25 | 0.33 | 0.71 | 0.07 | 0.08 | 0.99 | 0.88 | 0.15 | 0.15 | 0.09 | 0.17 | 0.35 | 0.82 | 0.07 | 0.52 | 0.12 | 0.15 | 0.94 | 0.22 | 0.03 | 0.8 | | | | | | | | | | | |

| | |
|---------------------------|-----------------------------|
| | 121 : José C. Paz |
| 101 : CABA | 122 : San Isidro |
| 102 : La Matanza | 123 : Vicente López |
| 103 : La Plata | 124 : Escobar |
| 104 : Lomas de Zamora | 125 : Ezeiza |
| 105 : Quilmes | 126 : Hurlingham |
| 106 : Almirante Brown | 127 : Ituzaingó |
| 107 : Merlo | 128 : San Fernando |
| 108 : Moreno | 129 : General Rodríguez |
| 109 : Florencio Varela | 130 : Zárate |
| 110 : Lanús | 131 : Luján |
| 111 : General San Martín | 132 : Campana |
| 112 : Tigre | 133 : Presidente Perón |
| 113 : Pilar | 134 : Berisso |
| 114 : Avellaneda | 135 : San Vicente |
| 115 : Tres de Febrero | 136 : Cañuelas |
| 116 : Berazategui | 137 : Marcos Paz |
| 117 : Malvinas Argentinas | 138 : Ensenada |
| 118 : Esteban Echeverría | 139 : Exaltación de la Cruz |
| 119 : Morón | 140 : Brandsen |
| 120 : San Miguel | 141 : General Las Heras. |

Dado $r > 0$ construimos un cubrimiento disjunto del conjunto de los 41 distritos de la siguiente forma:

$$A_{101}(r) = B_{d_1}(101, r)$$

$$A_{102}(r) = B_{d_1}(102, r) \setminus B_{d_1}(101, r)$$

$$A_{103}(r) = B_{d_1}(103, r) \setminus (B_{d_1}(101, r) \cup B_{d_1}(102, r))$$

$$\vdots$$

$$A_j(r) = B_{d_1}(j, r) \setminus \left(\bigcup_{l=101}^{j-1} B_{d_1}(l, r) \right); \quad j \leq 141.$$

Claro que varios A_j pueden ser vacíos. El algoritmo toma la forma

$$A_{101}(r) = \{101 \leq l \leq 141 : d_1(101, l) < r\}$$

$$A_{102}(r) = \{101 \leq l \leq 141 : d_1(102, l) < r \wedge d_1(101, l) \geq r\}$$

$$\vdots$$

$$A_j(r) = \{101 \leq l \leq 141 : d_1(j, l) < r \wedge d_1(m, l) \geq r \quad \forall m : 101 \leq m \leq j - 1\}; \quad j \leq 141.$$

Luego se eliminan los que son vacíos y se obtiene un cubrimiento disjunto de $\{1, \dots, 41\}$ dado por $\{A_j(r) : j \in I(r)\}$ donde $I(r) \subseteq \{1, \dots, 41\}$. Es decir $\{1, \dots, 41\} = \bigcup_{j \in I(r)} A_j(r)$.

Para cada r entre 0 y 1 y para cada $j \in I(r)$ elegimos la ciudad con mayor población dentro de $A_j(r)$. Es decir tomamos en $A_j(r)$ un elemento $k(j, r) \in I(r)$ tal que el número de habitantes de la ciudad $k(j, r)$ es mayor que el de todas las otras ciudades que conforman el clúster $A_j(r)$. Con esto construimos un nuevo grafo $\mathcal{G}_2(r)$ para cada r dado por $\mathcal{G}_2(r) = (\mathcal{V}_2(r), d_1)$ con $\mathcal{V}_2(r) = \{k(j, r) : j \in I(r)\}$ y d_1 como antes, es la métrica difusiva para $t = 100$. Con el algoritmo para la distancia de Hausdorff (Sección 4.2 Capítulo 4) estimamos $d_{GH}((\mathcal{V}_1, d_1), (\mathcal{V}_2(r), d_1)) \leq d_{1,H}(\mathcal{V}_1, \mathcal{V}_2(r))$ para $r = \frac{i}{50}$, $i = 1, \dots, 50$ y obtenemos la grafica de la Figura 6.6.7.

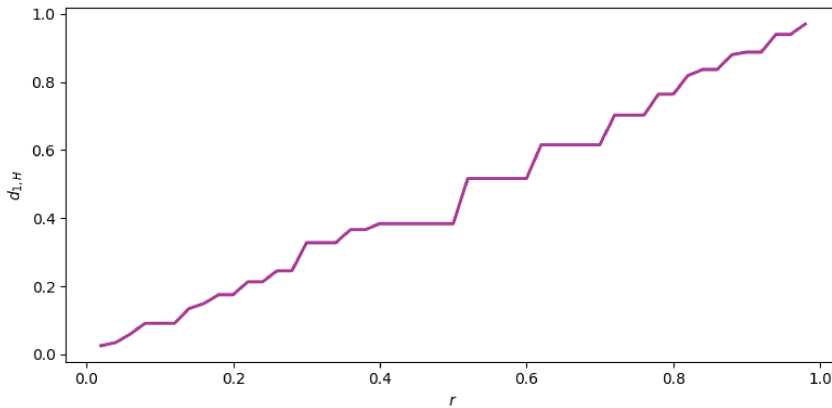


FIGURA 6.6.7. Cota superior para la evolución de $d_{GH}((\mathcal{V}_1, d_1), (\mathcal{V}_2(r), d_1))$ con $0 < r < 1$.

De alguna manera, al menos empíricamente resulta que la aplicación que a r entre 0 y 1 le asigna la distancia de Hausdorff con la métrica difusiva entre \mathcal{V}_1 y $\mathcal{V}_2(r)$ es de clase Lipschitz-1.

Con el objeto de ilustrar cuáles son los clústers generados por este algoritmo y los correspondientes $\mathcal{V}_2(r)$, mostramos algún detalle del caso $r = 0.3$ en la siguiente Figura 6.6.8.

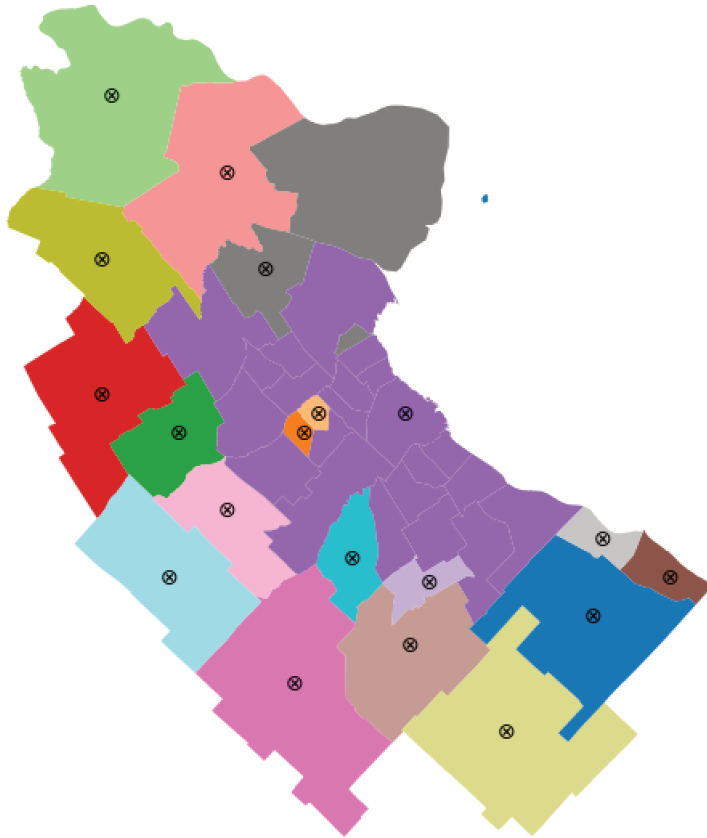


FIGURA 6.6.8. $\mathcal{V}_2(0.3)$ es 0.33 denso en \mathcal{V}_1 con d_1 .

Podemos visualizar los 19 clústers con diferentes colores y el símbolo \otimes marca el representante de cada clúster. Resulta que $\mathcal{V}_2(0.3) = \{\text{Zárate, Campana, Exaltación de la Cruz, Escobar, Luján, General Rodríguez, General Las Heras, Marcos Paz, Cañuelas, Ezeiza, La Plata, Presidente Perón, Brandsen, CABA, Ensenada, Berisso, Ituzaingó, San Vicente, Malvinas Argentinas}\}$, la distancia de Gromov-Hausdorff de este conjunto con \mathcal{V}_1 respecto a d_1 resulta acotada por 0.328.

6.7. Casi distancia de Gromov-Hausdorff entre espacios casi-métricos

Recordamos del Capítulo 4 que la extensión a espacios casi-métricos de la métrica de Hausdorff produce una casi-métrica entre los subconjuntos de aquel espacio. Esto permite extender la definición de Gromov a estas estructuras que, si bien son metrizables, proveen cierta flexibilidad adicional en la modelización de situaciones particulares.

Con el esquema provisto por la Proposición 6.2.1 podemos definir una distancia (casi-distancia) de Gromov-Hausdorff entre dos espacios casi-métricos (X_1, d_1) y (X_2, d_2) como $d_{GH}((X_1, d_1), (X_2, d_2)) = \inf\{d_{H,\delta}(X_1, X_2) : \text{donde } d_{H,\delta} \text{ es la casi-distancia de Hausdorff generada por una casi-métrica } \delta \text{ en } X_1 \sqcup X_2 \text{ que extiende tanto a } d_1 \text{ como a } d_2\}$.

El carácter casi-métrico de d_{GH} se deduce del carácter casi-métrico de $d_{H,\delta}$ que probamos en el Capítulo 4.

Observemos que el Teorema de metrización de Macías y Segovia [MS79] puede aplicarse en cualquier instancia casi-métrica. En particular puede aplicarse para metrizar esta última d_{GH} .

TEOREMA 6.7.1. *En las condiciones precedentes, existe una métrica ρ_{GH} y un número $\theta \geq 1$ tal que las desigualdades*

$$c_1 d_{GH}((X_1, d_1), (X_2, d_2)) \leq \rho_{GH}^\theta((X_1, d_1), (X_2, d_2)) \leq c_2 d_{GH}((X_1, d_1), (X_2, d_2))$$

valen para algún par de constantes $0 < c_1 \leq c_2 < \infty$.

Distancias entre medidas y distribuciones de probabilidad en espacios métricos

7.1. Introducción

En consonancia con la amplitud y universalidad de usos y aplicaciones de la teoría general de la medida, los abordajes a diversas ideas de definir y calcular distancias entre medidas y distribuciones, provienen de campos muy diversos de la matemática. Estas investigaciones están hoy amplificadas por las aplicaciones de sus ideas básicas en Teoría de la Información, aprendizaje automático y ciencia de datos. Los aportes intradisciplinarios más relevantes provienen de la Estadística Matemática, de la Teoría de Probabilidad, de la Teoría de la Información, del Análisis Matemático y de la Teoría Geométrica de la Medida. Este capítulo está dedicado a introducir la métrica de Kantorovich-Rubinstein entre probabilidades que será de utilidad en el desarrollo posterior de la tesis, su relación con la métrica de Wasserstein y a la descripción breve de otras nociones usuales de distancias entre medidas, algunas de las cuales no alcanzan el rango matemático de métricas, pero resultan útiles en aplicaciones.

7.2. Las métricas de Kantorovich-Rubinstein y de Wasserstein

Sea X un conjunto no vacío. Un álgebra de conjuntos en X es una colección no vacía \mathcal{A} de subconjuntos de X que es cerrado bajo uniones finitas y complementos, es decir

- i) Si $E_1, \dots, E_n \in \mathcal{A}$, entonces $\bigcup_{j=1}^n E_j \in \mathcal{A}$,
- ii) si $E \in \mathcal{A}$, entonces $E^c \in \mathcal{A}$.

Una σ -**álgebra** es un álgebra que es cerrada bajo uniones numerables. Es sencillo verificar que la intersección de cualquier familia de σ -álgebras en X es una σ -álgebra. Denotamos con $\mathcal{P}(X)$ al conjunto de todos los subconjuntos de X . Si $\mathcal{E} \subset \mathcal{P}(X)$, entonces la intersección de todas las σ -álgebras que contienen a \mathcal{E} . $\mathcal{M}(\mathcal{E})$ se llama la σ -álgebra **generada** por \mathcal{E} .

Si (X, d) es un espacio métrico, o más generalmente un espacio topológico la σ -álgebra generada por la familia de conjuntos abiertos en X se llama la **σ -álgebra de Borel** en X y se denota por $\mathcal{B}(X)$, que es la menor σ -álgebra de partes de X que contiene a todos los abiertos. Sea X un conjunto y $\mathcal{A} \subseteq \mathcal{P}(X)$ una σ -álgebra, el par (X, \mathcal{A}) se llama espacio medible.

Una medida positiva en X es una función $\mu : \mathcal{A} \rightarrow [0, \infty]$, con \mathcal{A} una σ -álgebra de las partes de X que cumple que la medida del vacío es cero y que la medida de uniones numerables disjuntas de elementos de la σ -álgebra es igual a la suma de las medidas de cada conjunto, es decir $\mu(\bigcup_k A_k) = \sum_k \mu(A_k)$, con $A \in \mathcal{A}$. Una probabilidad es una medida positiva tal que todo el espacio tiene medida unitaria. Usaremos $(\Omega, \mathcal{F}, \mathcal{P})$ para denotar un espacio de probabilidad. El conjunto Ω es el espacio muestral, \mathcal{F} es la σ -álgebra de eventos y \mathcal{P} es una probabilidad en \mathcal{F} .

La metrización de Kantorovich-Rubinstein de medidas probabilísticas, requiere que el conjunto básico sobre el que estarán definidas las medidas de probabilidad sea métrico o casi-métrico. Esto se debe a que se requiere una regularidad de Lipschitz de las funciones test. El concepto que definiremos a continuación es, de todos, el que mejor describe la convergencia en distribución de probabilidades o la convergencia vaga de las variables aleatorias que éstas describen.

DEFINICIÓN 7.2.1. Sea (Ω, d) un espacio métrico compacto con la σ -álgebra de Borel inducida por la topología que d genera en Ω . Sea $\mathfrak{M}(\Omega, d)$ el convexo de todas las medidas de Borel probabilísticas en Ω . La función

$$d_{KR}(\mu, \nu) = \sup \left\{ \left| \int_{\Omega} \varphi d\mu - \int_{\Omega} \varphi d\nu \right|, \phi : \Omega \rightarrow \mathbb{R} : |\varphi(\omega_1) - \varphi(\omega_2)| \leq d(\omega_1, \omega_2) \right\}$$

se denomina **métrica de Kantorovich-Rubinstein** en $\mathfrak{M}(\Omega, d)$.

Observemos que el supremo se toma sobre todas las funciones $\varphi : \Omega \rightarrow \mathbb{R}$ que son de clase Lipschitz 1 con constante menor o igual que 1.

Resumimos las propiedades más importantes de d_{KR} en el siguiente enunciado. Las pruebas pueden verse en muchas de las referencias bibliográficas en teoría de probabilidades. En particular en Billingley [Bil71].

TEOREMA 7.2.2. *Sea (Ω, d) un espacio métrico compacto. Entonces*

- a) d_{KR} es una métrica en $\mathfrak{M}(\Omega, d)$;
- b) $d_{KR}(\mu_n, \mu) \rightarrow 0$, $n \rightarrow \infty$ si y sólo si μ_n converge débilmente a μ ($\mu_n \rightharpoonup \mu$), es decir $\int_{\Omega} \varphi d\mu_n \rightarrow \int_{\Omega} \varphi d\mu$ para toda $\varphi : \Omega \rightarrow \mathbb{R}$ continua y acotada;
- c) (Prokhorov) el nuevo espacio métrico $(\mathfrak{M}(\Omega, d), d_{KR})$ también es compacto.

Introducimos ahora brevemente la métrica Wasserstein entre medidas de probabilidad en espacios métricos. Este enfoque de metrización de clases de medidas probabilísticas constituye la base teórica del transporte óptimo. El libro de Villani [Vil09] es la bibliografía básica más relevante desde el punto de vista teórico. Una referencia central para los métodos computacionales de transporte óptimo es el libro de Peyré y Cuturi [PC19]. En la base de la teoría está la noción de acoplamiento de medidas probabilísticas que procedemos a definir.

DEFINICIÓN 7.2.3. Sean (Ω_1, d_1) y (Ω_2, d_2) dos espacios métricos. Sean \mathcal{B}_1 y \mathcal{B}_2 las respectivas σ -álgebras de Borel. Sea $\Omega = \Omega_1 \times \Omega_2$ que resulta también métrico con $d((x_1, x_2), (y_1, y_2)) = \max\{d_1(x_1, y_1), d_2(x_2, y_2)\}$. Sea \mathcal{B} la σ -álgebra de Borel en Ω generada por d . Sean \mathfrak{M}_1 , \mathfrak{M}_2 y \mathfrak{M} las clases de medidas probabilísticas Borelianas en (Ω_1, d_1) , (Ω_2, d_2) y (Ω, d) respectivamente. Dadas $\mu_1 \in \mathfrak{M}_1$ y $\mu_2 \in \mathfrak{M}_2$ un **acoplamiento** entre μ_1 y μ_2 es cualquier $\pi \in \mathfrak{M}$ cuyas marginales son μ_1 y μ_2 . Precisamente

$$\pi_1(A) := \pi(A \times \Omega_2) = \mu_1(A), \quad A \in \mathcal{B}_1;$$

$$\pi_2(B) := \pi(\Omega_1 \times B) = \mu_2(B), \quad B \in \mathcal{B}_2.$$

Denotamos con $\mathfrak{A}(\mu_1, \mu_2)$ la familia de todos los acoplamientos entre μ_1 y μ_2 ,

$$\mathfrak{A}(\mu_1, \mu_2) = \{\pi \in \mathfrak{M} : \pi_1 = \mu_1 \text{ y } \pi_2 = \mu_2\}.$$

Algunas observaciones importantes relacionadas a las estructuras involucradas, incluyen el hecho de que el espacio (Ω, d) será polaco y/o compacto cuando (Ω_i, d_i) lo sean para $i = 1, 2$. Notemos también que para μ_1 y μ_2 dadas el conjunto $\mathfrak{A}(\mu_1, \mu_2)$ es convexo. Por otra parte, como $\pi = \mu_1 \times \mu_2$ tiene a μ_1 y μ_2 como primera y segunda marginales, si tiene que $\mathfrak{A}(\mu_1, \mu_2) \neq \emptyset$. El producto $\pi = \mu_1 \times \mu_2$ corresponde a independencia de las variables aleatorias, cuyas distribuciones son μ_1 y μ_2 , y que toman valores en los respectivos espacios métricos (Ω_1, d_1) y (Ω_2, d_2) . Suele decirse que $\mu_1 \times \mu_2$ es el acoplamiento

trivial. Con estas breves consideraciones, procedemos a definir la distancia de Wasserstein de orden p .

DEFINICIÓN 7.2.4. Sea $1 \leq p < \infty$, la **métrica o distancia de Wasserstein** de orden p entre $\mu_1 \in \mathfrak{M}_1$ y $\mu_2 \in \mathfrak{M}_2$ está dada por

$$\begin{aligned} W_p(\mu_1, \mu_2) &= \inf_{\pi \in \mathfrak{A}(\mu_1, \mu_2)} \left(\iint_{\Omega} d^p(x_1, x_2) d\pi(x_1, x_2) \right)^{\frac{1}{p}} \\ &= \inf_{\pi \in \mathfrak{A}(\mu_1, \mu_2)} \|d\|_{L^p(\Omega, d\pi)} \end{aligned}$$

Cuando $p = 1$,

$$W_1(\mu_1, \mu_2) = \inf_{\pi \in \mathfrak{A}(\mu_1, \mu_2)} \iint_{\Omega} d(x_1, x_2) d\pi(x_1, x_2).$$

En [Vil09] se demuestra que $W_1(\mu_1, \mu_2) = d_{KR}(\mu_1, \mu_2)$. Esta equivalencia resulta útil para obtener cotas superiores e inferiores de d_{KR} en casos particulares.

7.3. La distancia variación total

Sea (Ω, \mathcal{A}) un espacio de medida. Es decir Ω es un conjunto y \mathcal{A} es una σ -álgebra de partes de Ω . Sea $\mathfrak{M}(\Omega, \mathcal{A})$ la familia de todas las medidas de probabilidad definidas en el espacio de medida (Ω, \mathcal{A}) . Dadas μ y ν en $\mathfrak{M}(\Omega, \mathcal{A})$ se define la **distancia variación total** de μ a ν por

$$\mathbb{V}(\mu, \nu) = \sup_{A \in \mathcal{A}} |\mu(A) - \nu(A)|$$

PROPOSICIÓN 7.3.1. $\mathfrak{M}(\Omega, \mathcal{A})$ es un espacio métrico.

Observemos que si $\Omega = \mathbb{R}$, $\mathcal{A} = \mathcal{B}$ la σ -álgebra de Borel en \mathbb{R} y si $d\mu = g(x)dx$; $g \geq 0$, $\int_{\mathbb{R}} gdx = 1$ y si $\nu = \delta_0$ es la masa unitaria concentrada en cero entonces tenemos que, para $A = \{0\}$, $\mu(A) = \int_A gdx = 0$ y $\nu(A) = 1$. Por consiguiente $\mathbb{V}(\mu, \nu) = 1$. Notemos también que esto ocurre aún si g es la normal $N(0, \sigma^2)$ con σ tan chica como se quiera. Es decir que la normalidad con varianza muy pequeña no converge en el sentido de \mathbb{V} al determinismo. Esta es una de las razones por las que \mathbb{V} no es apta para la convergencia vaga de variables aleatorias y resulta $d_{KR} = W_1$ más adecuada. Cuando las medidas de probabilidad en $\mathfrak{M}(\Omega, \mathcal{A})$ tienen una medida dominante común la distancia variación total es la mitad de la norma L^1 de la diferencia de las densidades. Más aún el supremo que define \mathbb{V} es un máximo. Precisamos todo esto en el siguiente enunciado. La notación es la usual en el contexto de Radon-Nikodyn [Tay97].

PROPOSICIÓN 7.3.2. *Sea (Ω, \mathcal{A}) un espacio de medida. Sean μ y $\nu \in \mathfrak{M}(\Omega, \mathcal{A})$. Supongamos que existe una medida λ tal que $\mu \ll \lambda$ y $\nu \ll \lambda$. Sean $f_\mu = \frac{d\mu}{d\lambda}$ y $f_\nu = \frac{d\nu}{d\lambda}$. Entonces*

$$\mathbb{V}(\mu, \nu) = \frac{1}{2} \int_{\Omega} |f_\mu - f_\nu| d\lambda.$$

DEMOSTRACIÓN. Notar primero que

$$\begin{aligned} 0 &= \mu(\Omega) - \nu(\Omega) \\ &= \int_{\Omega} (f_\mu - f_\nu) d\lambda \\ &= \int_{\{f_\mu = f_\nu\}} (f_\mu - f_\nu) d\lambda + \int_{\{f_\mu > f_\nu\}} (f_\mu - f_\nu) d\lambda + \int_{\{f_\mu < f_\nu\}} (f_\mu - f_\nu) d\lambda \\ &= \int_{\{f_\mu > f_\nu\}} (f_\mu - f_\nu) d\lambda - \int_{\{f_\mu < f_\nu\}} (f_\nu - f_\mu) d\lambda. \end{aligned}$$

Entonces

$$\int_{\{f_\mu > f_\nu\}} (f_\mu - f_\nu) d\lambda = \int_{\{f_\mu < f_\nu\}} (f_\nu - f_\mu) d\lambda.$$

Por lo tanto

$$\begin{aligned} \int_{\Omega} |f_\mu - f_\nu| d\lambda &= \int_{\{f_\mu > f_\nu\}} (f_\mu - f_\nu) d\lambda + \int_{\{f_\mu < f_\nu\}} (f_\nu - f_\mu) d\lambda \\ &= 2 \int_{\{f_\mu > f_\nu\}} (f_\mu - f_\nu) d\lambda. \end{aligned}$$

Por otra parte

$$\begin{aligned} \int_{\{f_\mu > f_\nu\}} (f_\mu - f_\nu) d\lambda &= \max_{A \in \mathcal{A}} \int_A (f_\mu - f_\nu) d\lambda \\ &= \max_{A \in \mathcal{A}} |\mu(A) - \nu(A)| \\ &= \mathbb{V}(\mu, \nu), \end{aligned}$$

de donde se obtiene el primer resultado de la tesis. \square

7.4. Distancia de Hellinger

Sea (Ω, \mathcal{A}) un espacio de medida. Sea μ , ν y λ como en la Proposición 7.3.2. Es decir μ y ν pertenecen a $\mathfrak{M}(\Omega, \mathcal{A})$ y λ es una dominante común para μ y ν . En otras palabras $\mu \ll \lambda$ y $\nu \ll \lambda$. Por el Teorema de Radon-Nikodym tenemos que μ y ν

tienen densidades f_μ y f_ν respectivamente con respecto a λ . Entonces $f_\mu \geq 0$, $f_\nu \geq 0$ y además $\int_\Omega f_\mu d\lambda = \mu(\Omega) = 1$, $\int_\Omega f_\nu d\lambda = \nu(\Omega) = 1$. Es decir f_μ y f_ν están en la esfera unitaria de $L^1(\Omega, d\lambda)$. Como el espacio $L^1(\Omega, d\lambda)$ no es de Hilbert, conviene a veces, llevar estas densidades la bola unitaria de $L^2(\Omega, d\lambda)$. En particular, las funciones $g_\mu = \sqrt{f_\mu}$ y $g_\nu = \sqrt{f_\nu}$ están bien definidas y tienen norma $L^2(\Omega, d\lambda)$ igual a uno. En efecto $\|g_\mu\|_{L^2}^2 = \int_\Omega |g_\mu|^2 d\lambda = \int_\Omega |f_\mu| d\lambda = 1$ y del mismo modo $\|g_\nu\|_{L^2}^2 = 1$. La **distancia de Hellinger** entre μ y ν está dada por la norma $L^2(\Omega, d\lambda)$ de la diferencia $g_\mu - g_\nu$. Más precisamente

$$\begin{aligned} H_\lambda(\mu, \nu) &= \frac{1}{\sqrt{2}} \|g_\mu - g_\nu\|_{L^2(\Omega, d\lambda)} \\ &= \frac{1}{\sqrt{2}} \left\| \sqrt{\frac{d\mu}{d\lambda}} - \sqrt{\frac{d\nu}{d\lambda}} \right\|_{L^2(\Omega, d\lambda)}. \end{aligned}$$

Sea $\mathfrak{M}(\Omega, \mathcal{A}, \lambda)$, con λ una medida positiva y σ -finita, la familia de las medidas de probabilidad que están dominadas por λ .

PROPOSICIÓN 7.4.1. H_λ es una métrica en $\mathfrak{M}(\Omega, \mathcal{A}, \lambda)$.

PROPOSICIÓN 7.4.2. Sean μ y ν en $\mathfrak{M}(\Omega, \mathcal{A}, \lambda)$. Entonces

$$H_\lambda^2(\mu, \nu) = 1 - \int_\Omega \sqrt{\frac{d\mu}{d\lambda} \cdot \frac{d\nu}{d\lambda}} d\lambda.$$

DEMOSTRACIÓN.

$$\begin{aligned} H_\lambda^2(\mu, \nu) &= \frac{1}{2} \int_\Omega \left| \sqrt{\frac{d\mu}{d\lambda}} - \sqrt{\frac{d\nu}{d\lambda}} \right|^2 d\lambda \\ &= \frac{1}{2} \int_\Omega \left(\frac{d\mu}{d\lambda} + \frac{d\nu}{d\lambda} - 2\sqrt{\frac{d\mu}{d\lambda} \cdot \frac{d\nu}{d\lambda}} \right) d\lambda \\ &= \frac{1}{2} \left(\mu(\Omega) + \nu(\Omega) - 2 \int_\Omega \sqrt{\frac{d\mu}{d\lambda} \cdot \frac{d\nu}{d\lambda}} d\lambda \right) \\ &= 1 - \int_\Omega \sqrt{\frac{d\mu}{d\lambda} \cdot \frac{d\nu}{d\lambda}} d\lambda \end{aligned}$$

□

Notar que como $\sqrt{\frac{d\mu}{d\lambda}}$ y $\sqrt{\frac{d\nu}{d\lambda}}$ son positivas, entonces son perpendiculares si y sólo si la medida λ de la intersección de sus soportes es nula.

COROLARIO 7.4.3. *Sea μ y ν en $\mathfrak{M}(\Omega, \mathcal{A}, \lambda)$. Entonces $0 \leq H_\lambda(\mu, \nu) \leq 1$ y $H_\lambda(\mu, \nu) = 1$ si y sólo si $\lambda(\{\frac{d\mu}{d\lambda} > 0\} \cap \{\frac{d\nu}{d\lambda} > 0\}) = 0$.*

7.5. La protométrica de Bhattacharyya

Sea λ una medida positiva y σ -finita en (Ω, \mathcal{A}) . Consideremos como en la sección 7.4 el conjunto convexo $\mathfrak{M}(\Omega, \mathcal{A}, \lambda)$ de todas las medidas de probabilidad μ en (Ω, \mathcal{A}) que están dominadas por λ ($\mu \ll \lambda$). Como ya vimos en 7.4, dada μ y ν en $\mathfrak{M}(\Omega, \mathcal{A}, \lambda)$ tenemos que $\sqrt{\frac{d\mu}{d\lambda}}$ y $\sqrt{\frac{d\nu}{d\lambda}}$ están ambas en $L^2(\Omega, d\lambda)$ y tienen norma $L^2(\Omega, d\lambda)$ igual a uno. Por consiguiente, si θ es el ángulo (de Hilbert) entre $\sqrt{\frac{d\mu}{d\lambda}}$ y $\sqrt{\frac{d\nu}{d\lambda}}$, tenemos que $\cos \theta = \int_\Omega \sqrt{\frac{d\mu}{d\lambda}} \cdot \sqrt{\frac{d\nu}{d\lambda}} d\lambda$. Pero como esta integral es mayor o igual que cero, el ángulo θ está restringido a $0 \leq \theta \leq \frac{\pi}{2}$. Por consiguiente, también el producto escalar $\cos \theta$ está entre 0 y 1. Notemos que la distancia de Hellinger, con esta nueva notación, entre μ y ν está dada por $H_\lambda(\mu, \nu) = 1 - \cos \theta$, donde θ depende de μ y ν . Como vimos en la sección precedente $1 - \cos \theta(\mu, \nu)$ es una métrica en $\mathfrak{M}(\Omega, \mathcal{A}, \lambda)$.

La **protométrica de Bhattacharyya** está dada por

$$B_\lambda(\mu, \nu) = -\log \cos \theta(\mu, \nu).$$

PROPOSICIÓN 7.5.1. *Las propiedades métricas de B_λ son las siguientes:*

- a) $B_\lambda(\mu, \mu) = 0$;
- b) $B_\lambda(\mu, \nu) = 0$ implica $\mu = \nu$;
- c) $B_\lambda(\mu, \nu) = B_\lambda(\nu, \mu)$.

7.6. La protométrica de Kullback-Leibler

Consideremos, como en secciones precedentes, un espacio de medida $(\Omega, \mathcal{A}, \lambda)$ con λ positiva (no negativa) y σ -finita.

Cuando el conjunto convexo $\mathfrak{M}(\Omega, \mathcal{A}, \lambda)$ de las medidas probabilísticas en (Ω, \mathcal{A}) que son absolutamente continuas con respecto a λ se considera como el conjunto de distribuciones de ciertas variables aleatorias, la **divergencia de Kullback-Leibler** tienen una interpretación en términos de la teoría de la información. La frase que sigue es una traducción libre de una en el libro de Cover y Thomas [CT06], cuando introduce el concepto de información mutua de dos variables aleatorias "... Es la reducción en la incertidumbre de una variable aleatoria debida al conocimiento de otra".

Sea μ y ν en $\mathfrak{M}(\Omega, \mathcal{A}, \lambda)$ la **protométrica de Kullback-Leibler** (o entropía relativa) entre μ y ν está dada por

$$\begin{aligned} KL_\lambda(\mu||\nu) &= \int_{\Omega} \frac{d\mu}{d\lambda} \log \frac{\frac{d\mu}{d\lambda}}{\frac{d\nu}{d\lambda}} d\lambda \\ &= \int_{\Omega} f_\mu \log \frac{f_\mu}{f_\nu} d\lambda, \end{aligned}$$

siempre que la última integral exista.

Notemos que en el integrando de la última expresión para $KL_\lambda(\mu||\nu)$ entendemos que si $f_\mu = 0$ en algunos puntos de Ω , entonces $f_\mu \log \frac{f_\mu}{f_\nu} = 0$ en esos mismos puntos. Claro que si $f_\nu(\omega) = 0$ y $f_\mu(\omega) > 0$, entonces $f_\mu(\omega) \log \frac{f_\mu(\omega)}{f_\nu(\omega)} = +\infty$.

Resulta claro desde ejemplos sencillos que KL_λ no es simétrica ni satisface la desigualdad triangular. Queda sólo un pequeño espacio para el apelativo de protométrica a KL_λ .

PROPOSICIÓN 7.6.1. Sean $\mu, \nu \in \mathfrak{M}(\Omega, \mathcal{A}, \lambda)$. Entonces

- a) $KL_\lambda(\mu||\nu) \geq 0$;
- b) $KL_\lambda(\mu||\nu) = 0$ si y sólo si $\mu = \nu$.

DEMOSTRACIÓN. Como $f_\mu d\lambda = d\mu$ es medida positiva y probabilística, podemos aplicar la desigualdad de Jensen. Es decir, si ϕ es convexa,

$$\begin{aligned} \phi \left(\int_{\Omega} g d\mu \right) &\leq \int_{\Omega} \phi(g) d\mu \\ &= \int_{\Omega} \phi(g) f_\mu d\lambda. \end{aligned}$$

Entonces por la desigualdad de Jensen,

$$\begin{aligned} 1 = \nu(\Omega) &= \int_{\Omega} f_\nu d\lambda \\ &= \int_{\Omega} \frac{f_\nu}{f_\mu} f_\mu d\lambda \\ &= \int_{\Omega} \exp \left(\log \frac{f_\nu}{f_\mu} \right) d\mu \\ &= \int_{\Omega} \exp \left(-\log \frac{f_\mu}{f_\nu} \right) d\mu \end{aligned}$$

$$\begin{aligned} &\geq \exp\left(-\int_{\Omega} f_{\mu} \log \frac{f_{\mu}}{f_{\nu}} d\lambda\right) \\ &= \exp(-KL_{\lambda}(\mu||\nu)). \end{aligned}$$

Es decir que

$$e^{KL_{\lambda}(\mu||\nu)} \geq 1.$$

Por consiguiente $KL_{\lambda}(\mu||\nu) \geq 0$. Por otra parte, la igualdad vale si y sólo si $\log \frac{f_{\mu}}{f_{\nu}} = 0$. Es decir, si y sólo si $f_{\mu} = f_{\nu}$ y, por consiguiente, si $\mu = \nu$. \square

Una simetrización de KL_{λ} usual y simple es $\frac{1}{2}(KL_{\lambda}(\mu||\nu) + KL_{\lambda}(\nu||\mu))$.

7.7. Metrización de Kantorovich-Rubinstein de una clase de medidas probabilísticas discretas

Sea $n \in \mathbb{N}$, dado. Sea

$$S = \left\{ \bar{a} = (a_1, \dots, a_n) \in \mathbb{R}^n : a_i \geq 0 \text{ y } \sum_{i=1}^n a_i = 1 \right\}$$

y sea $\mathcal{V} = \{1, 2, \dots, n\}$. Para cada $\bar{a} \in S$, se tiene que la medida $\mu_{\bar{a}} : \mathcal{P}(\mathcal{V}) \rightarrow [0, 1]$ dada por $\mu_{\bar{a}}(E) = \sum_{i \in E} a_i$, $E \subset \mathcal{V}$ es una probabilidad en $\mathcal{P}(\mathcal{V})$. Sea d una métrica en \mathcal{V} . Entonces está bien definida la distancia de Kantorovich-Rubinstein, con respecto a la métrica d en \mathcal{V} , entre cualquier par de probabilidades $\mu_{\bar{a}}$ y $\mu_{\bar{b}}$ con $\bar{a}, \bar{b} \in S$. De esta manera tenemos un espacio métrico dado por $(\{(\mathcal{V}, d, \mu_{\bar{a}}) : \bar{a} \in S\}, d_{KR})$ y

$$d_{KR}(\mu_{\bar{a}}, \mu_{\bar{b}}) = \sup \left\{ \left| \int_{\mathcal{V}} \psi d\mu_{\bar{a}} - \int_{\mathcal{V}} \psi d\mu_{\bar{b}} \right| : |\psi(i) - \psi(j)| \leq d(i, j) \right\}.$$

PROPOSICIÓN 7.7.1. *La función $\gamma : S \rightarrow (\{(\mathcal{V}, d, \mu_{\bar{a}}) : \bar{a} \in S\}, d_{KR})$ dada por $\gamma(\bar{a}) = (\mathcal{V}, d, \mu_{\bar{a}})$ es continua si S hereda su topología de \mathbb{R}^n y en $(\{(\mathcal{V}, d, \mu_{\bar{a}}) : \bar{a} \in S\}, d_{KR})$ consideramos la topología generada por d_{KR} .*

DEMOSTRACIÓN. Sea $\bar{a}^k \rightarrow \bar{a}$, $k \rightarrow \infty$ con \bar{a}^k y \bar{a} todos en S . Entonces, si tomamos $\psi : (\mathcal{V}, d) \rightarrow (\mathbb{R}, |\cdot|)$, Lipschitz-1 con constante 1, tenemos que

$$\left| \int_{\mathcal{V}} \psi d\mu_{\bar{a}^k} - \int_{\mathcal{V}} \psi d\mu_{\bar{a}} \right| = \left| \sum_{j=1}^n \psi(j) a_j^k - \sum_{j=1}^n \psi(j) a_j \right|$$

$$\begin{aligned}
&= \left| \sum_{j=1}^n \psi(j)(a^k - a_j) \right| \\
&= \left| \sum_{j=1}^n (\psi(j) - \psi(1))(a_j^k - a_j) \right| \\
&\leq \sum_{j=1}^n |\psi(j) - \psi(1)| |a_j^k - a_j| \\
&\leq \sum_{j=1}^n d(j, 1) |a_j^k - a_j| \\
&\leq \sqrt{n} \cdot \text{diam}\mathcal{V} \|\bar{a}^k - \bar{a}\|,
\end{aligned}$$

que tiende a cero cuando $k \rightarrow \infty$. En la tercera igualdad hemos usado que $\sum_{j=1}^n (a_j^k - a_j) = 0$. \square

COROLARIO 7.7.2. *Sea $S = \{\bar{a} = (a_1, \dots, a_n) \in \mathbb{R}^n : a_i \geq 0 \text{ y } \sum_{i=1}^n a_i = 1\}$ y sea γ la función con dominio en S que a cada \bar{a} asigna el grafo métrico $(\mathcal{V}, d, \mu_{\bar{a}})$ con atributos dados por las componentes de \bar{a} . La función γ es Lipschitz con norma euclídea en \mathbb{R}^n en el dominio S y con la métrica de Kantorovich-Rubinstein d_{KR} en la imagen. Precisamente*

$$d_{KR}(\mu_{\bar{a}}, \mu_{\bar{b}}) \leq \sqrt{n} \text{diam}_d \mathcal{V} \|\bar{a} - \bar{b}\|,$$

con $\bar{a}, \bar{b} \in S$ y $\|\bar{a} - \bar{b}\|$ la norma euclídea.

DEMOSTRACIÓN. Inspección de la demostración de la Proposición 7.7.1. \square

7.8. Cálculo de algunas distancias entre distintos atributos en AMBA

Ya ilustramos en las secciones anteriores algunas diferencias teóricas esenciales entre distintas distancias entre medidas de probabilidad. En particular observamos que la de Kantorovich-Rubinstein-Wasserstein (1) es la que metriza la convergencia en distribuciones de variables aleatorias y en esto se basa su robustez. En esta breve sección exponemos los resultados del cálculo concreto en Python de las distancias mutuas entre los cuatros atributos en los distritos de AMBA que consideramos en nuestro análisis. Lo hacemos sólo para Kantorovich-Rubinstein, asociada a la métrica difusiva d_{100} , variación total y Hellinger con respecto a λ uniforme. La medida μ_1 es la uniforme en $\mathcal{V} = \{1, \dots, 41\}$ que representa los 41 distritos de AMBA, es decir $\mu_1(\{j\}) = \frac{1}{41}$ para todo j . Como $\mu_2(\{j\})$

tomamos la normalización de la población del distrito j . Como $\mu_3(\{j\})$ se toma la normalización de la densidad (superficial) de población del distrito j . Finalmente $\mu_4(\{j\})$ es una medida normalizada de la cantidad de infectados en julio de 2020. Los datos han sido tomados del trabajo de [AAGM22]. Explícitamente

$$\mu_1 \sim \left(\frac{1}{41}, \dots, \frac{1}{41} \right);$$

$$\begin{aligned} \mu_2 \sim & (0.0357936, 0.00578839, 0.0063182, 0.0152683, 0.0131109, 0.0065662, 0.0108293, \\ & 0.0277042, 0.00717137, 0.0324249, 0.0226366, 0.0104683, 0.00461876, 0.0160215, \\ & 0.0213332, 0.001857, 0.00376637, 0.0222016, 0.0309518, 0.0254558, 0.0184031, \\ & 0.0427359, 0.00397857, 0.0362991, 0.00634011, 0.0174921, 0.0277144, 0.00766766, \\ & 0.0218946, 0.184104, 0.00369825, 0.00218754, 0.00104226, 0.0115876, 0.136549, \\ & 0.0388071, 0.0215463, 0.0190729, 0.039793, 0.0182043, 0.0205954); \end{aligned}$$

$$\begin{aligned} \mu_3 \sim & (0.035909334, 0.005520003, 0.000812504, 0.001658514, 0.003534812, 0.003023856, \\ & 0.029208269, 0.080270388, 0.001289955, 0.019123166, 0.008661127, 0.001356966, \\ & 0.000762246, 0.068786437, 0.05320646, 0.000175903, 0.000351806, 0.007681096, \\ & 0.018779736, 0.06238692, 0.044562085, 0.007488441, 0.000887891, 0.011860886, \\ & 0.004188166, 0.050609797, 0.008535482, 0.000778999, 0.006483281, 0.11924546, \\ & 0.004305435, 0.000376935, 0.000167527, 0.040114253, 0.045634256, 0.057788313, \\ & 0.042811432, 0.048013134, 0.015479461, 0.025531059, 0.062638209); \end{aligned}$$

$\mu_4 \sim (0.0259413, 0.00169143, 0.000753787, 0.00658185, 0.00625092, 0.00194882, 0.00415502$
 $0.0259597, 0.00163627, 0.00902706, 0.00796073, 0.00490881, 0.0012318, 0.00814458$
 $0.032468, 0.000294161, 0.000717017, 0.0124283, 0.014322, 0.0240109, 0.00766657$
 $0.0109575, 0.00145242, 0.0113987, 0.00367701, 0.0118584, 0.0124467, 0.00152596$
 $0.0130166, 0.535777, 0.00121341, 0.000257391, 0.000478011, 0.00654508, 0.0741469$
 $0.0308869, 0.00845713, 0.0106633, 0.0356854, 0.0101669, 0.0212899).$

Para las métricas \mathbb{V} , H y d_{KR} entre ellas obtenemos los resultados que se muestran en el Cuadro 1.

| Atributos | $\mathbb{V}(\mu_i, \mu_j)$ | $H(\mu_i, \mu_j)$ | $d_{KR}(\mu_i, \mu_j)$ |
|----------------|----------------------------|-------------------|------------------------|
| μ_1, μ_2 | 0.3657 | 0.3548 | 0.8758 |
| μ_1, μ_3 | 0.4765 | 0.4303 | 0.8293 |
| μ_1, μ_4 | 0.5902 | 0.5557 | 0.7909 |
| μ_2, μ_3 | 0.3996 | 0.3221 | 0.4289 |
| μ_2, μ_4 | 0.3635 | 0.2934 | 0.3643 |
| μ_3, μ_4 | 0.4980 | 0.4024 | 0.3054 |

CUADRO 1. Distancias \mathbb{V} , H y d_{KR} mutuas entre μ_1 , μ_2 , μ_3 y μ_4 .

Capítulo 8

Una casi-métrica de tipo Gromov-Lipschitz en clases de espacios métricos con medidas probabilísticas.

El espacio $(\mathcal{M}^l(X_0, d_0), \mathfrak{d}_{GL}^{KR})$

8.1. Introducción

Los espacios métricos con medidas constituyen las extensiones naturales de los espacios Euclídeos equipados con la medida de Lebesgue que son los ambientes propicios para desarrollar y aplicar algunas técnicas del análisis armónico. Una estructura clásica en este sentido es la de espacio de tipo homogéneo, espacios casi-métricos con medidas duplicantes. Casos importantes, por sus aplicaciones, de tales contextos son los grafos en los que se introducen métricas y medidas y, atendiendo al carácter aleatorio que a veces asumen estas estructuras, es conveniente y aún relevante en aplicaciones, tener buenas nociones de distancias entre grafos que atiendan a la doble estructura que soportan: métrica y medida. Hemos introducido en los capítulos anteriores nociones varias de distancias entre espacios métricos y entre medidas probabilísticas.

En [ACI09] se introduce una métrica de este tipo bajo la hipótesis adicional de que los espacios cuyas distancias se pretende calcular son subespacios de un mismo ambiente compacto. Usando las ideas de metrización de Gromov extendemos esta construcción para espacios métricos probabilísticos generales y nos interesamos especialmente, contemplando las aplicaciones en capítulos siguientes, en grafos métricos no dirigidos y con atributos dados.

Precisamente, en la Sección 8.2, introducimos el enfoque Gromov-Lipschitz con métricas generales en las medidas y probamos sus propiedades básicas. Pero nos enfocamos especialmente en el caso de Kantorovich-Rubinstein que en la Sección 8.3 se aplicará al caso de grafos métricos con atributos, y cuando en el Capítulo 12 consideremos variables aleatorias valuadas en espacios métricos con medidas probabilísticas. En la Sección 8.4 aplicamos este concepto de distancia para el análisis de datos de SUBE en AMBA con

distintos atributos en los diferentes distritos de AMBA: población, densidad (con respecto al área) de la población, infectados de COVID-19 en julio de 2020 y el atributo uniforme en el que cada uno de los 41 distritos pesa $\frac{1}{41}$.

8.2. El espacio casi-métrico $(\mathcal{M}^l(X_0, d_0), \mathfrak{d}_{GL}^{KR})$

Como en el Capítulo 5, consideramos aquí dos espacios métricos entre los cuales existe alguna transformación bi-Lipschitz con el agregado adicional de una medida de Borel probabilística en cada uno de ellos. Más precisamente, sean (X, d, μ) e (Y, δ, ν) dados. Sea, como antes, $\Lambda = \{f : X \rightarrow Y \text{ bi-Lipschitz}\}$, para cada $f \in \Lambda$ es razonable definir las medidas probabilísticas $\tilde{\mu}_f$ y $\tilde{\nu}_f$ que se obtienen como “push forward” de f^{-1} y de f respectivamente. Precisamos, si A es un Boreliano en X , definimos $\tilde{\mu}_f(A) = \nu(f(A)) = \nu((f^{-1})^{-1}(A))$ y para B un Boreliano en Y definimos $\tilde{\nu}_f(B) = \mu(f^{-1}(B))$. Es claro que $\tilde{\mu}_f$ y $\tilde{\nu}_f$ son medidas de probabilidad, $\tilde{\mu}_f$ en X y $\tilde{\nu}_f$ en Y respectivamente. Por consiguiente todas las métricas y protométricas introducidas en el Capítulo 7 pueden aplicarse para calcular distancias entre μ y $\tilde{\mu}_f$ y ν y $\tilde{\nu}_f$ en X e Y respectivamente. Sean ρ_X y ρ_Y algunas de las distancias entre medidas probabilísticas como las consideradas en el Capítulo 7. Definimos la “distancia” de Gromov-Lipschitz respecto a ρ_X y ρ_Y entre (X, d, μ) e (Y, δ, ν) como

$$\mathfrak{d}_{GL}^{\rho_X \rho_Y}((X, d, \mu), (Y, \delta, \nu)) = \inf_{f \in \Lambda} \{|\log \text{dil}(f)| + |\log \text{dil}(f^{-1})| + \rho_X(\mu, \tilde{\mu}_f) + \rho_Y(\nu, \tilde{\nu}_f)\}.$$

La siguiente proposición reúne algunas propiedades elementales de $\mathfrak{d}_{GL}^{\rho_X \rho_Y}$.

PROPOSICIÓN 8.2.1. *Consideremos $\mathfrak{d}_{GL}^{\rho_X \rho_Y}$ sobre la clase de todos los espacios métricos compactos con medidas de Borel probabilísticas. Entonces*

- a) $\mathfrak{d}_{GL}^{\rho_X \rho_Y}((X, d, \mu), (X, d, \mu)) = 0$;
- b) $\mathfrak{d}_{GL}^{\rho_X \rho_Y}((X, d, \mu), (Y, \delta, \nu)) = \mathfrak{d}_{GL}^{\rho_Y \rho_X}((Y, \delta, \nu), (X, d, \mu))$;
- c) $\mathfrak{d}_{GL}^{\rho_X \rho_Y} < \infty$ si y sólo si $\Lambda \neq \emptyset$;
- d) $\mathfrak{d}_{GL}^{\rho_X \rho_Y}((X, d, \mu), (Y, \delta, \nu)) = 0$ implica que existe $f : (X, d) \rightarrow (Y, \delta)$ que es una isometría y además $\mu = \nu \circ f$.

DEMOSTRACIÓN. Las propiedades de la a) hasta la c) son simples. También lo es d) si usamos el inciso e) de la Proposición 5.2.4 puesto que por hipótesis (X, d) e (Y, δ) son compactos. Como $\mathfrak{d}_{GL}^{\rho_X \rho_Y}((X, d, \mu), (Y, \delta, \nu)) = 0$, entonces $d_{GL}((X, d), (Y, \delta)) = 0$. Por lo

tanto (X, d) e (Y, δ) son isométricos. Sea f una isometría entre ellos entonces puesto que $\tilde{\mu}_f = \nu \circ f$ y $\rho(\mu, \tilde{\mu}_f) = 0$ resulta que $\mu = \tilde{\mu}_f = \nu \circ f$ y también $\nu = \tilde{\nu}_f = \mu \circ f^{-1}$. \square

Para obtener una desigualdad de tipo triangular restringimos la familia de espacios y consideramos que ρ_X y ρ_Y son métricas de Kantorovich-Rubinstein en cada espacio. De esta forma podemos obtener todavía casi-métricas sobre familias especiales de espacios métricos compactos con medidas de Borel probabilísticas. Sea (X_0, d_0) un espacio métrico compacto dado y fijo. Sea l un número natural con $l \geq 2$. Sea $\mathcal{M}^l(X_0, d_0)$ el conjunto de todos los (X, d) espacios métricos tales que existe una transformación bi-Lipschitz $f : (X_0, d_0) \rightarrow (X, d)$ con $\text{dil}(f) \leq l$ y $\text{dil}(f^{-1}) \leq l$. En otras palabras

$$\frac{1}{l} \leq \frac{d(f(x_1^0), f(x_2^0))}{d_0(x_1^0, x_2^0)} \leq l$$

para todo par x_1^0 y x_2^0 en X_0 distintos.

Los siguientes resultados serán usados en la prueba de una propiedad casi-triangular para d_{GL}^{KR} , la distancia de Gromov-Lipschitz cuando usamos la métrica de Kantorovich-Rubinstein en cada espacio métrico, en $\mathcal{M}^l(X_0, d_0) := \{(X, d, \mu) : (X, d) \in \mathcal{M}^l(X_0, d_0) \text{ y } \mu \text{ es una medida de probabilidad en } (X, d)\}$. Observamos que si $(X, d, \mu) \in \mathcal{M}^l(X_0, d_0)$, entonces (X, d) es compacto ya que (X, d) es homeomorfo a (X_0, d_0) , pues son bi-Lipschitz equivalentes.

LEMA 8.2.2. *Si (X_i, d_i) con $i = 1, 2$ están en $\mathcal{M}^l(X_0, d_0)$ entonces existe $\varphi : (X_1, d_1) \rightarrow (X_2, d_2)$ bi-Lipschitz con $\text{dil}(\varphi) \leq l^2$ y $\text{dil}(\varphi^{-1}) \leq l^2$.*

DEMOSTRACIÓN. Como (X_i, d_i) están en $\mathcal{M}^l(X_0, d_0)$ con $i = 1, 2$ entonces existen dos transformaciones bi-Lipschitz $f : (X_0, d_0) \rightarrow (X_1, d_1)$ y $g : (X_0, d_0) \rightarrow (X_2, d_2)$ tales que

$$\frac{1}{l} \leq \frac{d_1(f(x), f(y))}{d_0(x, y)} \leq l,$$

y

$$\frac{1}{l} \leq \frac{d_2(g(x), g(y))}{d_0(x, y)} \leq l.$$

Sea $\varphi = g \circ f^{-1} : (X_1, d_1) \rightarrow (X_2, d_2)$, entonces para $x \neq y$ con x, y en X_1

$$\frac{d_2(\varphi(x), \varphi(y))}{d_1(x, y)} = \frac{d_2(g(f^{-1}(x)), g(f^{-1}(y)))}{d_1(x, y)}.$$

Tomando $x = f(u)$ e $y = f(v)$ obtenemos

$$\begin{aligned} \frac{d_2(g(f^{-1}(x)), g(f^{-1}(y)))}{d_1(x, y)} &= \frac{d_2(g(u), g(v))}{d_1(f(u), f(v))} \\ &= \frac{d_2(g(u), g(v))}{d_0(u, v)} \cdot \frac{d_0(u, v)}{d_1(f(u), f(v))} \\ &\leq l^2. \end{aligned}$$

□

LEMA 8.2.3. Si $f : (X_1, d_1) \rightarrow (X_2, d_2)$ es bi-Lipschitz y μ_2 es una medida probabilística en (X_2, d_2) entonces $\mu_2 \circ f$ es una medida probabilística en (X_1, d_1) y además

$$\int_{X_1} \eta d(\mu_2 \circ f) = \int_{X_2} (\eta \circ f^{-1}) d\mu_2,$$

para η medible, siempre que alguna de las dos integrales exista.

DEMOSTRACIÓN. Con los argumentos usuales en teoría de la medida, vemos que basta probar la fórmula para $\eta = \chi_E$ con E Boreliano en X_1 . Cuando $\eta = \chi_E$,

$$\begin{aligned} \int_{X_1} \eta d(\mu_2 \circ f) &= \int_E d(\mu_2 \circ f) \\ &= (\mu_2 \circ f)(E) \\ &= \mu_2(f(E)) \\ &= \int_{f(E)} d\mu_2 \\ &= \int_{X_2} \chi_{f(E)} d\mu_2 \\ &= \int_{X_2} (\chi_E \circ f^{-1}) d\mu_2. \end{aligned}$$

□

LEMA 8.2.4. Sean $(X_i, d_i) \in \mathcal{M}^l(X_0, d_0)$; $i = 1, 2$, $f : X_1 \rightarrow X_2$ bi-Lipschitz y $\eta : X_1 \rightarrow \mathbb{R}$ una función real Lipschitz 1, con respecto a d_1 , y con constante 1. Entonces

$$\begin{array}{ccc} X_1 & \xrightarrow{f} & X_2 \\ \eta \downarrow & \searrow \psi & \\ \mathbb{R} & & \end{array}$$

$\psi = \frac{\eta \circ f^{-1}}{\text{dil}(f^{-1})} : X_2 \rightarrow \mathbb{R}$ es de clase Lipschitz 1, con respecto a d_2 , con constante 1.

DEMOSTRACIÓN. Para x_2 y x'_2 en X_2 , con $x_2 \neq x'_2$, tenemos

$$\begin{aligned} \frac{|\psi(x_2) - \psi(x'_2)|}{d_2(x_2, x'_2)} &= \frac{1}{\text{dil}(f^{-1})} \frac{|\eta(f^{-1}(x_2)) - \eta(f^{-1}(x'_2))|}{d_2(x_2, x'_2)} \\ &= \frac{1}{\text{dil}(f^{-1})} \cdot \frac{|\eta(f^{-1}(x_2)) - \eta(f^{-1}(x'_2))|}{d_1(f^{-1}(x_2), f^{-1}(x'_2))} \cdot \frac{d_1(f^{-1}(x_2), f^{-1}(x'_2))}{d_2(x_2, x'_2)} \\ &\leq 1. \end{aligned}$$

□

PROPOSICIÓN 8.2.5. $(\mathcal{M}^l(X_0, d_0), \mathfrak{d}_{GL}^{KR})$ es un espacio casi-métrico con constante triangular menor o igual que l^2 .

DEMOSTRACIÓN. Sean (X_i, d_i, μ_i) con $i = 1, 2, 3$ tres espacios en $\mathcal{M}^l(X_0, d_0)$. Sea $\epsilon > 0$ dado. Sean $f : X_1 \rightarrow X_2$ y $g : X_2 \rightarrow X_3$ tales que ambas son bi-Lipschitz y cumplen (8.1)

$$|\log \text{dil}(f)| + |\log \text{dil}(f^{-1})| + \rho_{KR}(\mu_1, \tilde{\mu}_f^1) + \rho_{KR}(\mu_2, \tilde{\mu}_f^2) < \frac{\epsilon}{2} + \mathfrak{d}_{GL}^{\rho_{KR}}((X_1, d_1, \mu_1), (X_2, d_2, \mu_2)),$$

con $\tilde{\mu}_f^1 = \mu_2 \circ f$ y $\tilde{\mu}_f^2 = \mu_1 \circ f^{-1}$;

(8.2)

$$|\log \text{dil}(g)| + |\log \text{dil}(g^{-1})| + \rho_{KR}(\mu_2, \tilde{\mu}_g^2) + \rho_{KR}(\mu_3, \tilde{\mu}_g^3) < \frac{\epsilon}{2} + \mathfrak{d}_{GL}^{\rho_{KR}}((X_2, d_2, \mu_2), (X_3, d_3, \mu_3)),$$

con $\tilde{\mu}_g^2 = \mu_3 \circ g$ y $\tilde{\mu}_g^3 = \mu_2 \circ g^{-1}$.

Sea $h : X_1 \rightarrow X_3$ dada por $h = g \circ f$ que como vimos en la Proposición (5.2.4 d)) resulta bi-Lipschitz con $|\log \text{dil}(g \circ f)| \leq |\log \text{dil}(f)| + |\log \text{dil}(g)|$ y $|\log \text{dil}(g \circ f)^{-1}| \leq |\log \text{dil}(f^{-1})| + |\log \text{dil}(g^{-1})|$. Sean $\tilde{\mu}_h^1 = \mu_3 \circ h$ y $\tilde{\mu}_h^3 = \mu_1 \circ h^{-1}$. Entonces $\tilde{\mu}_h^1 = \mu_3 \circ h = (\mu_3 \circ g) \circ f = \tilde{\mu}_g^2 \circ f$ y

$$\begin{aligned} \rho_{KR}(\mu_1, \tilde{\mu}_h^1) &\leq \rho_{KR}(\mu_1, \tilde{\mu}_f^1) + \rho_{KR}(\tilde{\mu}_f^1, \tilde{\mu}_h^1) \\ &= \rho_{KR}(\mu_1, \tilde{\mu}_f^1) + \rho_{KR}(\mu_2 \circ f, \tilde{\mu}_g^2 \circ f). \end{aligned}$$

Mientras que el primer término de la derecha de esta desigualdad está en el miembro izquierdo de (8.1), el segundo término de la derecha en la última estimación se parece, pero no coincide exactamente, con el tercer sumando del miembro izquierdo de (8.2). Procedemos a estimar $\rho_{KR}(\mu_2 \circ f, \tilde{\mu}_g^2 \circ f)$ por definición de la métrica de Kantorovich-Rubinstein.

Sea $\eta : X_1 \rightarrow \mathbb{R}$ de clase Lipschitz 1 con constante 1, entonces por el Lema 8.2.3

$$\begin{aligned} \left| \int_{X_1} \eta \cdot d(\mu_2 \circ f) - \int_{X_1} \eta \cdot d(\tilde{\mu}_g^2 \circ f) \right| &= \left| \int_{X_2} (\eta \circ f^{-1}) d\mu_2 - \int_{X_2} (\eta \circ f^{-1}) d\tilde{\mu}_g^2 \right| \\ &= |\text{dil}(f^{-1})| \left| \int_{X_2} \frac{\eta \circ f^{-1}}{|\text{dil}(f^{-1})|} d\mu_2 - \int_{X_2} \frac{\eta \circ f^{-1}}{|\text{dil}(f^{-1})|} d\tilde{\mu}_g^2 \right|. \end{aligned}$$

Como por el Lema 8.2.4 $\psi = \frac{\eta \circ f^{-1}}{|\text{dil}(f^{-1})|}$, $\psi : X_2 \rightarrow \mathbb{R}$ es Lipschitz 1 con constante 1 tenemos

$$\begin{aligned} \rho_{KR}(\mu_2 \circ f, \tilde{\mu}_g^2 \circ f) &\leq |\text{dil}(f^{-1})| \rho_{KR}(\mu_2, \tilde{\mu}_g^2) \\ &\leq l^2 \rho_{KR}(\mu_2, \tilde{\mu}_g^2), \end{aligned}$$

en la última desigualdad hemos usado el Lema 8.2.2. Procedemos de manera análoga con $\rho_{KR}(\mu_3, \tilde{\mu}_h^3)$, obtenemos

$$\begin{aligned} \rho_{KR}(\mu_3, \tilde{\mu}_h^3) &\leq \rho_{KR}(\mu_3, \tilde{\mu}_g^3) + \rho_{KR}(\tilde{\mu}_g^3, \tilde{\mu}_h^3) \\ &\leq \rho_{KR}(\mu_3, \tilde{\mu}_g^3) + \rho_{KR}(\mu_2 \circ g^{-1}, \tilde{\mu}_f^2 \circ g^{-1}) \\ &\leq \rho_{KR}(\mu_3, \tilde{\mu}_g^3) + l^2 \rho_{KR}(\mu_2, \tilde{\mu}_f^2). \end{aligned}$$

Por consiguiente,

$$\begin{aligned} \mathfrak{d}_{GL}^{\rho_{KR}}((X_1, d_1, \mu_1), (X_3, d_3, \mu_3)) &\leq |\log \text{dil}(h)| + |\log \text{dil}(h^{-1})| + \rho_{KR}(\mu_1, \tilde{\mu}_h^1) + \rho_{KR}(\mu_3, \tilde{\mu}_h^3) \\ &\leq |\log \text{dil}(f)| + |\log \text{dil}(g)| + |\log \text{dil}(f^{-1})| + |\log \text{dil}(g^{-1})| \\ &\quad + \rho_{KR}(\mu_1, \tilde{\mu}_f^1) + l^2 \rho_{KR}(\mu_2, \tilde{\mu}_g^2) \\ &\quad + \rho_{KR}(\mu_3, \tilde{\mu}_g^3) + l^2 \rho_{KR}(\mu_2, \tilde{\mu}_f^2) \\ &\leq l^2 \{ \epsilon + \mathfrak{d}_{GL}^{\rho_{KR}}((X_1, d_1, \mu_1), (X_2, d_2, \mu_2)) \\ &\quad + \mathfrak{d}_{GL}^{\rho_{KR}}((X_2, d_2, \mu_2), (X_3, d_3, \mu_3)) \}. \end{aligned}$$

Como ϵ es arbitrario tenemos el resultado. \square

La Proposición 8.2.5 provee una casi-métrica en el espacio $\mathcal{M}^l(X_0, d_0)$. Pero, el Teorema de Macías y Segovia [MS79], afirma que dada una casi-métrica d en un conjunto con constante triangular K , existen $\beta \geq 1$, δ una métrica en X y dos constantes $0 < c_1 \leq c_2 < \infty$ tales que las desigualdades

$$c_1 d(x_1, x_2) \leq \delta^\beta(x_1, x_2) \leq c_2 d(x_1, x_2)$$

valen para todo par de puntos x_1 y x_2 en el conjunto. Las constantes dependen de l . De esta manera, si preferimos métricas, en vez de casi-métricas en $\mathcal{M}^l(X_0, d_0)$ podemos tomar una métrica δ_{GL}^{KR} tal que $d_{GL}^{KR} \sim (\delta_{GL}^{KR})^\beta$ y $(\mathcal{M}^l(X_0, d_0), \delta_{GL}^{KR})$ es espacio métrico.

8.3. Propiedades de continuidad de grafos métricos con atributos con la distancia de Gromov-Lipschitz-Kantorovich

Con los resultados de las Secciones 5.3 y 7.7 consideramos ahora la doble estructura de métrica y medida en un caso particular que usaremos más adelante. Recordemos que (Ξ, ρ) es un espacio métrico dado y que $\mathcal{V} = \{1, 2, \dots, n\}$, $n \in \mathbb{N}$. El conjunto $\{\xi_j^0 : j = 1, \dots, n\}$ es $4R$ -disperso en Ξ .

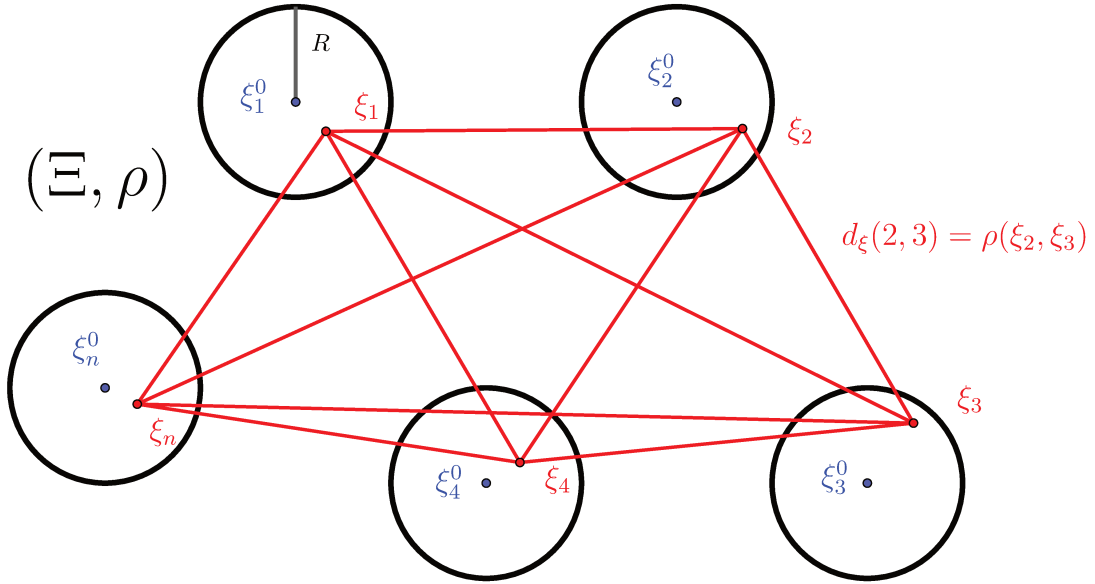


FIGURA 8.3.1. Disposición de los puntos $\{\xi_j^0\}$ y de los $\{\xi_j\}$ en (Ξ, ρ) .

Con la notación de la Sección 8.2 tenemos $X_0 = \mathcal{V}$ y $d_0(i, j) = \rho(\xi_i^0, \xi_j^0)$. Como en la Sección 5.3 del Capítulo 5 consideramos el espacio métrico (Ξ^n, ρ^n) con $\rho^n(\bar{\xi}, \bar{\eta}) = \max\{\rho(\xi_i, \eta_i) : i = 1, \dots, n\}$. Para cada $\bar{\xi} \in B_{\rho^n}(\bar{\xi}^0, R)$ tenemos una métrica $d_{\bar{\xi}}$ en \mathcal{V} dada por $d_{\bar{\xi}}(i, j) = \rho(\xi_i, \xi_j)$ para $i, j \in \mathcal{V}$. Aquí $\bar{\xi}^0 = (\xi_1^0, \dots, \xi_n^0)$.

PROPOSICIÓN 8.3.1. Si $\bar{\xi} \in B_{\rho^n}(\bar{\xi}^0, R)$ entonces $f : (\mathcal{V}, d_0) \rightarrow (\mathcal{V}, d_{\bar{\xi}})$ dada por la identidad $f(i) = i$, satisface

$$\frac{R}{D} \leq \frac{d_{\bar{\xi}}(f(i), f(j))}{d_0(i, j)} \leq \frac{D}{R},$$

donde $D = \sup\{\rho(\xi, \xi') : \xi, \xi' \in \bigcup_{i=1}^n B_\rho(\xi_i^0, R)\}$. Por consiguiente $\Sigma^l(\mathcal{V}, d_0) = \{(\mathcal{V}, d_{\bar{\xi}}) : \bar{\xi} \in B_{\rho^n}(\bar{\xi}^0, R)\} \subset \mathcal{M}^l(\mathcal{V}, d_0)$ para todo $l \geq \frac{D}{R}$.

DEMOSTRACIÓN. Sea $f : \mathcal{V} \rightarrow \mathcal{V}$ la identidad. Es decir $f(j) = j$, $j \in \mathcal{V}$. Entonces $d_{\bar{\xi}}(f(i), f(j)) = d_{\bar{\xi}}(i, j) = \rho(\xi_i, \xi_j)$. Por otra parte $d_0(i, j) = \rho(\xi_i^0, \xi_j^0)$. Para $i \neq j$ tenemos que $\rho(\xi_i, \xi_i^0) < R$, $\rho(\xi_j, \xi_j^0) < R$ y $\rho(\xi_i^0, \xi_j^0) \geq 4R$. Por consiguiente $R < \rho(\xi_i, \xi_j)$. Supongamos que no. Entonces

$$\rho(\xi_i^0, \xi_j^0) \leq \rho(\xi_i^0, \xi_i) + \rho(\xi_i, \xi_j) + \rho(\xi_j, \xi_j^0) < 4R,$$

que contradice la $4R$ dispersión de $\{\xi_i^0 : i = 1, \dots, n\}$. Entonces, para $i \neq j$,

$$\frac{R}{D} \leq \frac{\rho(\xi_i, \xi_j)}{\rho(\xi_i^0, \xi_j^0)} = \frac{d_{\bar{\xi}}(i, j)}{d_0(i, j)} \leq \frac{D}{R}.$$

□

Sea, como en la Sección 7.7, $S = \{\bar{a} = (a_1, \dots, a_n) \in \mathbb{R}^n : a_i \geq 0 \text{ y } \sum_{i=1}^n a_i = 1\}$, que interpretamos como el conjunto de todas las posibles medidas probabilísticas soportadas en \mathcal{V} . Por consiguiente, tenemos una subfamilia de $\mathcal{M}^l(\mathcal{V}, d_0)$ indexada en $B_{\rho^n}(\bar{\xi}^0, R) \times S$ dada por

$$\Sigma^l(\mathcal{V}, d_0) = \{(\mathcal{V}, d_{\bar{\xi}}, \mu_{\bar{a}}) : \bar{\xi} \in B_{\rho^n}(\bar{\xi}^0, R), \bar{a} \in S\},$$

que por las Proposiciones 8.3.1 y 8.2.5 resulta un espacio casi-métrico con d_{GL}^{KR} .

El resultado que reúne los de las Secciones 5.3 y 7.7 y que usaremos en el Capítulo 12 es el siguiente

TEOREMA 8.3.2. *Sea (Ξ, ρ) un espacio métrico. Sea $\mathcal{V} = \{1, 2, \dots, n\}$. Sea $\{\xi_i^0 : i \in \mathcal{V}\}$ un conjunto $4R$ -disperso en (Ξ, ρ) , $R > 0$. Para $\bar{\xi} = (\xi_1, \dots, \xi_n)$ sea $d_{\bar{\xi}} : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}_0^+$, $d_{\bar{\xi}}(i, j) = \rho(\xi_i, \xi_j)$. Sea $S = \{\bar{a} = (a_1, \dots, a_n) \in \mathbb{R}^n : a_i > 0, \sum_{i=1}^n a_i = 1\}$ y para cada $\bar{a} \in S$, $\mu_{\bar{a}}$ la medida de probabilidad en \mathcal{V} tal que $\mu_{\bar{a}}(\{i\}) = a_i$. Sea Γ la aplicación que a cada par $(\bar{\xi}, \bar{a})$ con $\bar{\xi} \in B_{\rho^n}(\bar{\xi}^0, R)$ y $\bar{a} \in S$ asigna el espacio métrico probabilístico dado por $(\mathcal{V}, d_{\bar{\xi}}, \mu_{\bar{a}})$ es de clase Lipschitz uno si en el dominio tomamos las métricas naturales definidas por ρ y $\|\cdot\|$ y en la imagen consideramos la casi-métrica de Gromov-Lipschitz-Kantorovich. Más aún la constante de Lipschitz puede estimarse uniformemente*

en términos de las constantes geométricas. Precisamente, la desigualdad

$$\mathfrak{d}_{GL}^{KR}((\mathcal{V}, d_{\bar{\xi}}, \mu_{\bar{a}}), (\mathcal{V}, d_{\bar{\eta}}, \mu_{\bar{b}})) \leq \Theta \left(\sum_{i=1}^n \rho^2(\xi_i, \eta_i) + \sum_{i=1}^n |a_i - b_i|^2 \right)^{\frac{1}{2}},$$

vale para toda elección de $\bar{\xi}$ y $\bar{\eta}$ tales que $\xi_i, \eta_i \in B_\rho(\xi_i^0, R)$ que satisfacen $\sum_{i=1}^n \rho^2(\xi_i, \eta_i) < R^2$ y para toda elección de \bar{a} y \bar{b} en S y donde $\Theta = 4 \max \left\{ \frac{2}{R}, \sqrt{n} \gamma + 2R \right\}$ con $\gamma = \text{diam}_{d_{\xi^0}} \mathcal{V} = \text{diam}_\rho \{\xi_1^0, \dots, \xi_n^0\}$.

DEMOSTRACIÓN. Para estimar superiormente $\mathfrak{d}_{GL}^{KR}((\mathcal{V}, d_{\bar{\xi}}, \mu_{\bar{a}}), (\mathcal{V}, d_{\bar{\eta}}, \mu_{\bar{b}}))$ con $\bar{a}, \bar{b} \in S$ y $\bar{\xi}, \bar{\eta} \in B_{\rho^n}(\xi^0, R)$, usamos la definición de \mathfrak{d}_{GL}^{KR} y la transformación bi-Lipschitz trivial dada por $f = \text{identidad}$. Es decir $f(j) = j$, $j = 1, \dots, n$. Para hacer las estimaciones usaremos el esquema del Teorema 5.3.5 y del Corolario 7.7.2. En efecto, observemos que, por definición de $\tilde{\mu}_f$ y $\tilde{\nu}_f$ en la Sección 8.2, cuando los espacios métricos probabilísticos son (X, d, μ) e (Y, δ, ν) , tenemos que $\tilde{\mu}_f = \nu \circ f$, $\tilde{\nu}_f = \mu \circ f^{-1}$. Ahora esos dos espacios son $(\mathcal{V}, d_{\bar{\xi}}, \mu_{\bar{a}})$ y $(\mathcal{V}, d_{\bar{\eta}}, \mu_{\bar{b}})$ y $f = \text{identidad}$ en \mathcal{V} . Es decir $\mu_{\bar{a}} = \mu$, $\mu_{\bar{b}} = \nu$ y $f = i_d$. Entonces $\tilde{\mu}_f = \mu_{\bar{b}} \circ i_d = \mu_{\bar{b}}$ y $\tilde{\nu}_f = \mu_{\bar{a}} \circ i_d^{-1} = \mu_{\bar{a}}$. Por consiguiente

$$\mathfrak{d}_{GL}^{KR}((\mathcal{V}, d_{\bar{\xi}}, \mu_{\bar{a}}), (\mathcal{V}, d_{\bar{\eta}}, \mu_{\bar{b}})) \leq |\log \text{dil}_{\bar{\xi}\bar{\eta}}(i_d)| + |\log \text{dil}_{\bar{\eta}\bar{\xi}}(i_d)| + \rho_{(\mathcal{V}, d_{\bar{\xi}})}^{KR}(\mu_{\bar{a}}, \mu_{\bar{b}}) + \rho_{(\mathcal{V}, d_{\bar{\eta}})}^{KR}(\mu_{\bar{a}}, \mu_{\bar{b}}),$$

donde

$$\begin{aligned} \text{dil}_{\bar{\xi}\bar{\eta}} &= \sup_{\substack{i, j \in \mathcal{V} \\ i \neq j}} \frac{d_{\bar{\xi}}(i, j)}{d_{\bar{\eta}}(i, j)} = \sup_{\substack{i, j \in \mathcal{V} \\ i \neq j}} \frac{\rho(\xi_i, \xi_j)}{\rho(\eta_i, \eta_j)} \\ \text{dil}_{\bar{\eta}\bar{\xi}} &= \sup_{\substack{i, j \in \mathcal{V} \\ i \neq j}} \frac{d_{\bar{\eta}}(i, j)}{d_{\bar{\xi}}(i, j)} = \sup_{\substack{i, j \in \mathcal{V} \\ i \neq j}} \frac{\rho(\eta_i, \eta_j)}{\rho(\xi_i, \xi_j)}. \end{aligned}$$

En la demostración del Teorema 5.3.5 probamos que

$$|\log \text{dil}_{\bar{\xi}\bar{\eta}}(i_d)| + |\log \text{dil}_{\bar{\eta}\bar{\xi}}(i_d)| \leq \frac{4}{R} \left(\sum_{i=1}^n \rho^2(\xi_i, \eta_i) \right)^{\frac{1}{2}}.$$

Por otra parte, usando el Corolario 7.7.2 y el hecho que los espacios métricos $(\mathcal{V}, d_{\bar{\xi}})$ y $(\mathcal{V}, d_{\bar{\eta}})$ son uniformemente acotados, tenemos que $\rho_{(\mathcal{V}, d_{\bar{\xi}})}^{KR}(\mu_{\bar{a}}, \mu_{\bar{b}}) + \rho_{(\mathcal{V}, d_{\bar{\eta}})}^{KR}(\mu_{\bar{a}}, \mu_{\bar{b}})$ están acotados por una constante veces $\|\bar{b} - \bar{a}\|$ de donde se sigue la tesis. Observamos que la constante Θ se obtiene tomando el máximo de $\frac{4}{R}$ y la constante de Corolario 7.7.2. \square

Recordemos (ver [Chu01]) que una medida ν positiva en (Ω, \mathcal{F}) es una subprobabilidad si $\nu(\Omega) \leq 1$. Observamos que tanto la definición \mathfrak{d}_{GL}^{KR} como el Teorema 8.3.2 siguen siendo válidos si en lugar de S tomamos $\tilde{S} = \{\bar{a} : a_i > 0, \sum_{i=1}^n a_i \leq 1\}$. Entonces, luego de normalizaciones adecuadas, el Teorema 8.3.2 también se extiende a $\tilde{S}_c = \{\bar{a} : a_i > 0, \sum_{i=1}^n a_i \leq c\}$ para $c > 0$.

Observamos también que aunque la estimación que provee el Teorema 8.3.2 resulta adecuada como una propiedad de continuidad, en general estará lejos de ser óptima, en particular porque en las estimaciones de la parte Gromov-Lipschitz de \mathfrak{d}_{GL}^{KR} hemos sólo usado el coeficiente de dilatación de la identidad. Esto significa que, en principio, $\mathfrak{d}_{GL}^{KR}((\mathcal{V}, d_{\bar{\xi}}, \mu_{\bar{a}}), (\mathcal{V}, d_{\bar{\eta}}, \mu_{\bar{b}}))$ puede ser bastante más chico que

$$\left(\sum_{i=1}^n \rho^2(\xi_i, \eta_i) + \sum_{i=1}^n |a_i - b_i|^2 \right)^{\frac{1}{2}}.$$

8.4. Aplicación del enfoque Lipschitz-Gromov al análisis de SUBE en AMBA

Procedemos ahora a usar nuevamente los datos SUBE con distintos atributos para estimar superiormente “las distancias” entre los espacios con medida definidos por $(\mathcal{V}, d_{100}, \mu_i)$ y $(\mathcal{V}, d_t, \mu_j)$ con $\mathcal{V} = \{1, \dots, 41\}$, d_t la métrica definida para t en un entorno de 100 y μ_i es cualquier par de los cuatro atributos en los vértices introducidos en la Sección 7.8 y que repasamos a continuación. En cada caso, puesto que la σ -álgebra es la de las partes de \mathcal{V} , nos basta definir $\mu_i(\{j\}) = \mu_{i,j}$ para $i = 1, 2, 3, 4$ y $j \in \mathcal{V}$. La medida μ_1 es la uniforme, es decir $\mu_1(\{j\}) = \frac{1}{41}$ para todo j . Para $\mu_2(\{j\})$ tomamos la normalización de la población del distrito j . Como $\mu_3(\{j\})$ se toma la normalización de la densidad (superficial) de población del distrito j . Finalmente $\mu_4(\{j\})$ es una medida normalizada de la cantidad de infectados en julio de 2020. Explícitamente

$$\mu_1 \sim \left(\frac{1}{41}, \dots, \frac{1}{41} \right);$$

$$\mu_2 \sim (0.0357936, 0.00578839, 0.0063182, 0.0152683, 0.0131109, 0.0065662, 0.0108293, \\ 0.0277042, 0.00717137, 0.0324249, 0.0226366, 0.0104683, 0.00461876, 0.0160215, \\ 0.0213332, 0.001857, 0.00376637, 0.0222016, 0.0309518, 0.0254558, 0.0184031, \\ 0.0427359, 0.00397857, 0.0362991, 0.00634011, 0.0174921, 0.0277144, 0.00766766, \\ 0.0218946, 0.184104, 0.00369825, 0.00218754, 0.00104226, 0.0115876, 0.136549, \\ 0.0388071, 0.0215463, 0.0190729, 0.039793, 0.0182043, 0.0205954);$$

$$\mu_3 \sim (0.035909334, 0.005520003, 0.000812504, 0.001658514, 0.003534812, 0.003023856, \\ 0.029208269, 0.080270388, 0.001289955, 0.019123166, 0.008661127, 0.001356966, \\ 0.000762246, 0.068786437, 0.05320646, 0.000175903, 0.000351806, 0.007681096, \\ 0.018779736, 0.06238692, 0.044562085, 0.007488441, 0.000887891, 0.011860886, \\ 0.004188166, 0.050609797, 0.008535482, 0.000778999, 0.006483281, 0.11924546, \\ 0.004305435, 0.000376935, 0.000167527, 0.040114253, 0.045634256, 0.057788313, \\ 0.042811432, 0.048013134, 0.015479461, 0.025531059, 0.062638209);$$

$$\mu_4 \sim (0.0259413, 0.00169143, 0.000753787, 0.00658185, 0.00625092, 0.00194882, 0.00415502 \\ 0.0259597, 0.00163627, 0.00902706, 0.00796073, 0.00490881, 0.0012318, 0.00814458 \\ 0.032468, 0.000294161, 0.000717017, 0.0124283, 0.014322, 0.0240109, 0.00766657 \\ 0.0109575, 0.00145242, 0.0113987, 0.00367701, 0.0118584, 0.0124467, 0.00152596 \\ 0.0130166, 0.535777, 0.00121341, 0.000257391, 0.000478011, 0.00654508, 0.0741469 \\ 0.0308869, 0.00845713, 0.0106633, 0.0356854, 0.0101669, 0.0212899).$$

Notar que para μ_2 los valores más grandes de $d_{GL}^{KR}((\mathcal{V}, d_{100}, \mu_2), (\mathcal{V}, d_t, \mu_2))$ se obtienen para los distritos más poblados, CABA con 0.184104 y la Matanza con 0.136549. Notar también que μ_4 refleja que CABA tiene la mayor cantidad de infectados en julio 2020. Los datos han sido tomados del trabajo de [\[AAGM22\]](#).

Con estas notaciones podemos estimar superiormente las distancias de Variación Total y de Hellinger entre cada par de μ_1, μ_2, μ_3 y μ_4 . Las figuras siguientes muestran cotas superiores $M_{ij}^{VT}(t)$ y $M_{ij}^H(t)$ para las distancias de Gromov-Lipschitz entre (d_{100}, μ_i) y de

(d_t, μ_j) para t en un entorno de 100 e i y j entre uno y cuatro para la métrica de Variación Total que muestra la Figura 8.4.2 y la métrica de Hellinger de la Figura 8.4.3 y en cada grafica se explicita una referencias de color.

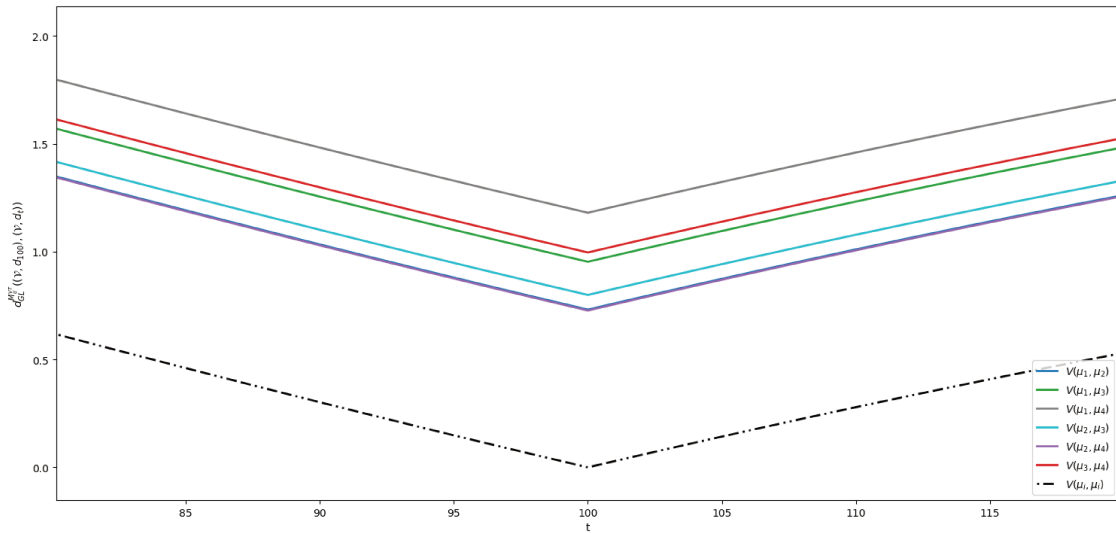


FIGURA 8.4.2. Gráfica de $d_{GL}^{MVT}((\mathcal{V}, d_{100}), (\mathcal{V}, d_t))$.

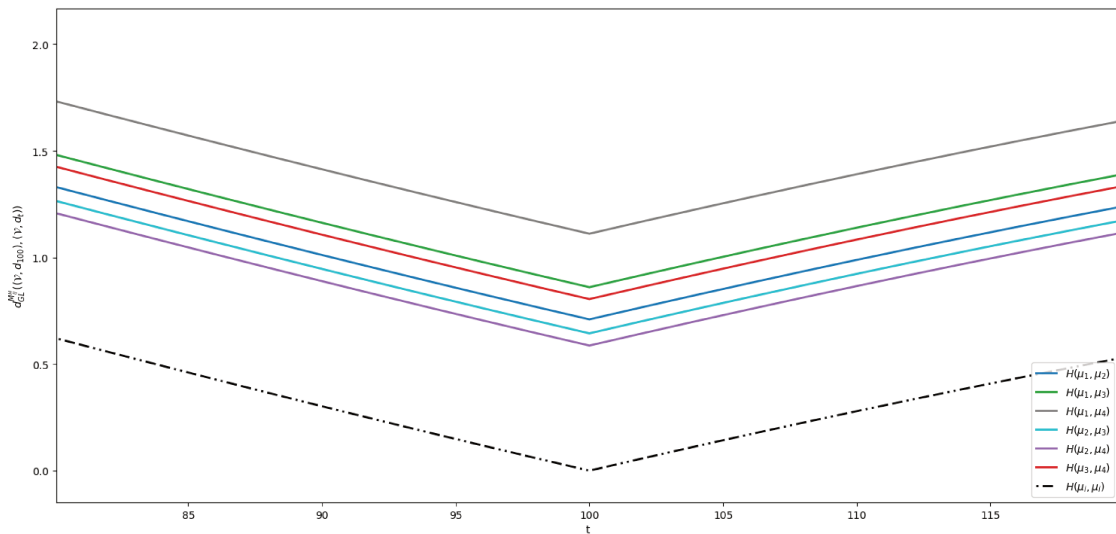


FIGURA 8.4.3. Gráfica de $d_{GL}^{MH}((\mathcal{V}, d_{100}), (\mathcal{V}, d_t))$.

Observamos finalmente que las secuencias de colores no preservan el mismo orden para Variación Total y para Hellinger, en efecto, mientras que para Variación Total tenemos

$$M_{ii}^{VT}(t) < M_{24}^{VT}(t) < M_{12}^{VT}(t) < M_{23}^{VT}(t) < M_{13}^{VT}(t) < M_{34}^{VT}(t) < M_{14}^{VT}(t),$$

para Hellinger tenemos

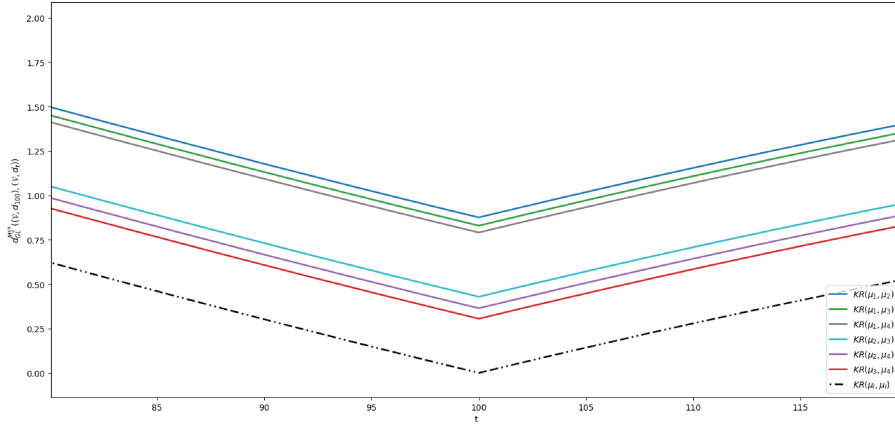
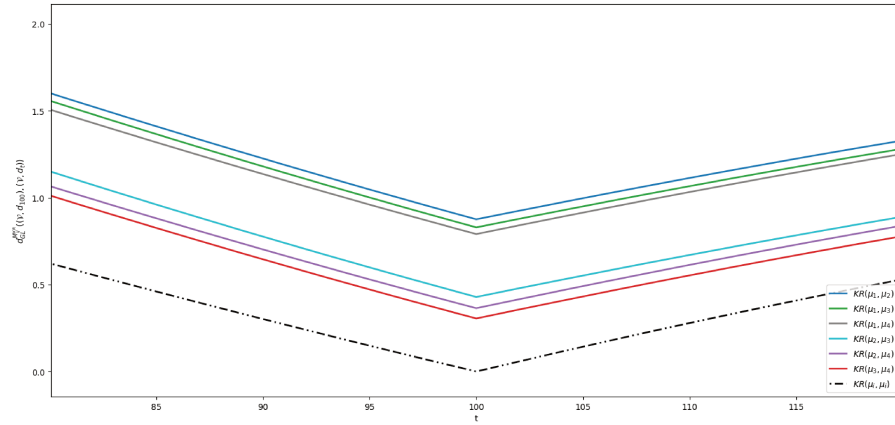
$$M_{ii}^H(t) < M_{24}^H(t) < M_{23}^H(t) < M_{12}^H(t) < M_{34}^H(t) < M_{13}^H(t) < M_{14}^H(t).$$

En lo que respecta a la métrica de Kantorovich-Rubinstein entre los atributos en el ejemplo de SUBE que, estamos considerando, la transformación bi-Lipschitz natural que nos provee la cota superior para la distancia entre las métricas difusivas, es la identidad. Pero en el enfoque Kantorovich-Rubinstein, la métrica se define como un supremo sobre las funciones Lipschitz 1 con constante unitaria, en principio sólo es sencillo obtener una cota inferior para d_{KR} . El enfoque dual provisto por la versión Wasserstein ($p = 1$) d_{W_1} de d_{KR} , nos permite obtener cotas superiores, posiblemente excesivas, pero simples usando acoplamientos triviales. En efecto si μ y ν son las medidas probabilísticas que determinan dos cualesquiera de los atributos que estamos considerando, la trivialidad del acoplamiento π , que corresponde a la independencia de las variables aleatorias distribuidas de acuerdo a μ y ν respectivamente, corresponde a $\pi = \mu \times \nu$. Ver C. Villani (Capítulo 1) [Vil09]. Con este punto de vista podemos entonces obtener, usando Python también las gráficas correspondientes a estimaciones superiores para $\rho_X^{KR} = \rho_X^{W_1}$. En esta estimación, si $\mu(\{i\}) = \mu_i$ y $\nu(\{j\}) = \nu_j$ con $i, j = 1, 2, \dots, 41$, tenemos que $\pi(\{(i, j)\}) = \pi_{ij} = \mu_i \cdot \nu_j$ y por consiguiente será necesario calcular

$$\sum_{m=1}^{41} \left(\sum_{l=1}^{41} d_t(l, m) \mu_l \right) \nu_m$$

que, en términos espectrales es

$$\sum_{m=1}^{41} \sum_{l=1}^{41} \left(\sum_{j=1}^{41} e^{-2t\lambda_j} |\varphi_j(l) - \varphi_j(m)|^2 \right)^{\frac{1}{2}} \mu_l \cdot \nu_m$$

FIGURA 8.4.4. Gráfica de $d_{GL}^{M^{KR}(100)}((\mathcal{V}, d_{100}), (\mathcal{V}, d_t))$.FIGURA 8.4.5. Gráfica de $d_{GL}^{M^{KR}(t)}((\mathcal{V}, d_{100}), (\mathcal{V}, d_t))$.

Observamos que las secuencias de colores preservan el mismo orden para la métrica de Kantorovich-Rubinstein en los casos de $t = 100$ y de t en un entorno de 100, en efecto, para $t = 100$ se puede visualizar en Figura 8.4.4 que tenemos la siguientes desigualdades $M_{ii}^{KR}(t) < M_{34}^{KR}(100) < M_{24}^{KR}(100) < M_{23}^{KR}(100) < M_{14}^{KR}(100) < M_{13}^{KR}(100) < M_{12}^{KR}(100)$, y para t en un entorno tenemos la misma desigualdad que se puede ver en la Figura 8.4.5

$$M_{ii}^{KR}(t) < M_{34}^{KR}(t) < M_{24}^{KR}(t) < M_{23}^{KR}(t) < M_{14}^{KR}(t) < M_{13}^{KR}(t) < M_{12}^{KR}(t).$$

Observamos también empíricamente en las gráficas de las Figuras 8.4.4 y 8.4.5 la formación de dos grupos de curvas. Las tres más altas (azul, verde y gris) corresponden

todas a la comparación de los atributos “físicos” dados por μ_2 , μ_3 y μ_4 con el más artificial dado por la distribución uniforme que es provista por μ_1 . En tanto que las tres inferiores (cian, púrpura y rojo) detectan alguna relación intrínseca razonable y esperada entre población μ_2 , densidad de población μ_3 e infectados μ_4 .

Capítulo 9

El enfoque Hausdorff-Gromov en la proto-metrización de espacios métricos de probabilidad

9.1. Introducción

En este breve capítulo recuperamos el enfoque de Hausdorff-Gromov para producir una plausible proto-métrica en clases de tipo \mathcal{M} de espacios métricos con medidas probabilísticas. No avanzaremos en ningún aspecto teórico de la idea y nos limitaremos a calcular la magnitud propuesta en algunos ejemplos sencillos y en nuestro persistente análisis de SUBE en AMBA.

9.2. La proto-métrica de Gromov-Hausdorff-Kantorovich

Aunque todo lo que sigue puede plantearse en el contexto general del Capítulo 6, para nuestras aplicaciones posteriores y con el objeto de obtener una metrización más sencilla, supondremos en esta sección que todos los espacios métricos que consideremos tendrán una copia isométrica dentro de un espacio métrico fijo al que denotaremos (\mathcal{Z}, ∂) .

Sea (\mathcal{Z}, ∂) un espacio métrico. Sea $\mathcal{M}(\mathcal{Z}, \partial)$ el conjunto de todos los espacios métricos (X, d) tales que existe una inmersión isométrica f de (X, d) en (\mathcal{Z}, ∂) . Es decir $f : X \rightarrow \mathcal{Z}$ y $\partial(f(x_1), f(x_2)) = d(x_1, x_2)$ para todo par de puntos x_1, x_2 en X . En forma compacta

$$\mathcal{M}(\mathcal{Z}, \partial) = \{(X, d) \text{ espacio métrico tal que existe } f : X \rightarrow \mathcal{Z}, \text{ inmersión isométrica}\}.$$

Dentro de esta clase de espacios métricos consideraremos la clase de espacios métricos probabilísticos dada por

$$\mathcal{M}(\mathcal{Z}, \partial) = \{(X, d, \mu) : (X, d) \in \mathcal{M}(\mathcal{Z}, \partial) \text{ y } (X, d) \text{ es compacto y } \mu \text{ es una probabilidad en } (X, d)\}.$$

Dado $(X, d, \mu) \in \mathcal{M}(\mathcal{Z}, \partial)$ y dada f isometría entre X y $f(X) \subset \mathcal{Z}$, consideramos el *push forward* de μ por f , que es la medida de Borel probabilística definida por

$$\mu_f = \mu \circ f^{-1},$$

que explícitamente significa que para cada Boreliano B en (\mathcal{Z}, ∂) , $\mu_f(B) = \mu(f^{-1}(B))$, ya que la continuidad de f asegura que $f^{-1}(B)$ es un Boreliano en (X, d) . Claro que $\mu_f(\mathcal{Z}) = \mu(f^{-1}(\mathcal{Z})) = \mu(X) = 1$, y por lo tanto μ_f , es una medida de probabilidad en (\mathcal{Z}, ∂) .

La conjunción de las dos teorías de Gromov-Hausdorff por una parte y de Kantorovich-Rubinstein por otra, nos permite proto-metrizar de una manera natural el conjunto $\mathcal{M}(\mathcal{Z}, \partial)$. En efecto, sean (X_i, d_i, μ_i) con $i = 1, 2$ dos elementos de $\mathcal{M}(\mathcal{Z}, \partial)$, definimos la **proto-métrica de Gromov-Hausdorff-Kantorovich-Rubinstein** por

$$\mathfrak{d}_{GH}^{KR}((X_1, d_1, \mu_1), (X_2, d_2, \mu_2)) = \inf_{\substack{f_1 \in \mathcal{I}(X_1, \mathcal{Z}) \\ f_2 \in \mathcal{I}(X_2, \mathcal{Z})}} \text{máx}\{d_{H, \partial}(f_1(X_1), f_2(X_2)); d_{K, \partial}(\mu_{f_1}^1, \mu_{f_2}^2)\},$$

donde $\mathcal{I}(X_i, \mathcal{Z})$ con $i = 1, 2$ es el conjunto de todas las inmersiones isométricas de (X_i, d_i) en (\mathcal{Z}, ∂) ; $d_{H, \partial}$ es la distancia de Hausdorff en el espacio (\mathcal{Z}, ∂) y $d_{K, \partial}$ es la métrica de Kantorovich-Rubinstein en las probabilidades Borelianas sobre (\mathcal{Z}, ∂) .

PROPOSICIÓN 9.2.1. *La función \mathfrak{d}_{GH}^{KR} definida en $\mathcal{M}(\mathcal{Z}, \partial)$ tiene las siguientes propiedades,*

- (1) $\mathfrak{d}_{GH}^{KR} : \mathcal{M}(\mathcal{Z}, \partial) \times \mathcal{M}(\mathcal{Z}, \partial) \rightarrow \mathbb{R}^+ \cup \{0\}$, es no negativa y finita;
- (2) \mathfrak{d}_{GH}^{KR} es simétrica;
- (3) \mathfrak{d}_{GH}^{KR} se anula en la diagonal de $\mathcal{M}(\mathcal{Z}, \partial) \times \mathcal{M}(\mathcal{Z}, \partial)$.

Consideremos algunos ejemplos sencillos.

EJEMPLO 9.2.1. Supongamos primero que tenemos dos espacios métricos de un solo punto $X = \{x\}$, $Y = \{y\}$. Claro que $d \equiv 0$ y $\delta \equiv 0$. Por otra parte, como μ y ν son probabilísticas, resulta que μ es la masa unitaria en x y ν es la masa unitaria en y . Sean $(Z, \partial) = (X, d)$, $f : X \rightarrow Z$, $f(x) = x$, $g : Y \rightarrow Z$, $g(y) = x$, con $\partial = d \equiv 0$. Entonces f y g son isometrías. También $f(X) = g(Y)$ por consiguiente $d_{H, \partial}(f(X), g(Y)) = 0$. Por otra parte $\mu_f = \mu$ y $\nu_g = \nu \circ g^{-1}$ es la masa unitaria en x . Es decir $\nu_g = \mu = \mu_f$. Entonces

$d_{K,\partial}(\mu_f, \nu_g) = 0$. Por consiguiente

$$\mathfrak{d}_{GH}^{KR}((X, d, \mu), (Y, \delta, \nu)) = 0.$$

EJEMPLO 9.2.2. Una ventaja de esta métrica comparada con Gromov-Lipschitz es que los cardinales de X y de Y pueden diferir y todavía pueden tener distancia finita. Supongamos que $X = \{x_0\}$ tiene sólo un punto. Sea (X, d, δ_{x_0}) , $d \equiv 0$ y $\mu = \delta_{x_0}$ la masa unitaria en x_0 . Sean $Y = \{y_1, y_2\}$, δ una métrica en Y y $\nu = \theta_1\delta_{y_1} + \theta_2\delta_{y_2}$, con $\theta_i \geq 0$, $\theta_1 + \theta_2 = 1$. Estimemos $\mathfrak{d}_{GH}^K((X, d, \mu), (Y, \delta, \nu))$. Sean $(Z, \partial) = (Y, \delta)$, $f : X \rightarrow Z$, $f(x_0) = y_1$ y $g : Y \rightarrow Z$, $g(y_1, y_2) = (y_1, y_2)$ la identidad. Es claro que f y g son inmersiones isométricas en (Z, ∂) . Por otra parte, $f(X) = \{f(x_0)\} = \{y_1\}$ y también $g(Y) = Y$. Como $(Z, \partial) = (Y, \delta)$, tenemos que

$$d_{H,\partial}(f(X), g(Y)) = d_{H,\delta}(\{y_1\}, Y).$$

Como $y_1 \in Y$, tenemos que cualquier ϵ -entorno de Y , que coincide con Y , contiene a $\{y_1\}$. Por otra parte, el ínfimo de los $\epsilon > 0$ para los cuales un ϵ -entorno de $\{y_1\}$ contiene a Y es $\delta(y_1, y_2)$. Por consiguiente $d_{H,\partial}(f(X), g(X)) = \delta(y_1, y_2)$. Ahora consideremos las probabilidades inducidas por μ y ν en (Z, ∂) . Tenemos que $\mu \circ f^{-1} = \delta_{y_1}$ y que $\nu \circ g^{-1} = \nu = \theta_1\delta_{y_1} + \theta_2\delta_{y_2}$. Calculemos en el espacio métrico $(Z, \partial) = (Y, \delta)$ la distancia de Kantorovich entre $\mu \circ f^{-1}$ y $\nu \circ g^{-1}$,

$$\begin{aligned} d_{K,\partial}(\mu \circ f^{-1}, \nu \circ g^{-1}) &= \sup \left| \int_Y \varphi d\mu \circ f^{-1} - \int_Y \varphi d\nu \circ g^{-1} \right| \\ &= \sup |\varphi(y_1) - \theta_1\varphi(y_1) - \theta_2\varphi(y_2)|, \end{aligned}$$

donde el supremo se toma sobre todas las funciones reales φ tales que $|\varphi(y_1) - \varphi(y_2)| \leq \delta(y_1, y_2) = \partial(y_1, y_2)$. Notemos que para una tal φ , tenemos

$$\begin{aligned} |(1 - \theta_1)\varphi(y_1) - \theta_2\varphi(y_2)| &= |\theta_2\varphi(y_1) - \theta_2\varphi(y_2)| \\ &= \theta_2|\varphi(y_1) - \varphi(y_2)| \\ &\leq \theta_2\delta(y_1, y_2) \\ &\leq \delta(y_1, y_2). \end{aligned}$$

Por consiguiente $d_{K,\delta}(\mu \circ f^{-1}, \nu \circ g^{-1}) \leq \delta(y_1, y_2)$ y entonces

$$d_{GH}^{KR}((X, d, \mu), (Y, \delta, \nu)) \leq \delta(y_1, y_2).$$

EJEMPLO 9.2.3. Sea $\mathcal{B} = \{B_i : i = 1, \dots, n\}$ una familia de discos (bolas) planos con distintos centros y posiblemente distintos radios. Digamos que $B_i = B(x_i, r_i)$ con $x_i = (x_i^1, x_i^2)$ y $r_i > 0$ $i = 1, \dots, n$ que satisface la “ley de conservación” $\sum_{i=1}^n r_i = 1$. Este esquema puede verse primero como un grafo métrico con atributos en sus vértices de la manera siguiente. El conjunto de vértices $\mathcal{V} = \{x_1, \dots, x_n\}$, es el conjunto de los centros, sus aristas \mathcal{A} son todos los pares de elementos de \mathcal{V} y la distancia entre ellos $d(x_i, x_j) = |x_i - x_j| = \sqrt{(x_i^1 - x_j^1)^2 + (x_i^2 - x_j^2)^2}$ es la que \mathcal{V} hereda de \mathbb{R}^2 . Los atributos r_i de los nodos x_i definen una medida de probabilidad μ en \mathcal{V} . De este modo (\mathcal{V}, d, μ) es un espacio métrico con medida. Estimemos la distancia de Hausdorff - Gromov - Kantorovich - Wasserstein entre dos de tales estructuras. Esta distancia puede luego entenderse como una métrica entre familias de bolas o entre “cubrimientos” en el plano. Sean $\mathcal{B} = \{B_i : i = 1, \dots, n\}$ y $\tilde{\mathcal{B}} = \{\tilde{B}_j : j = 1, \dots, m\}$ dos sistemas de discos planos como el descrito en \mathbb{R}^2 .

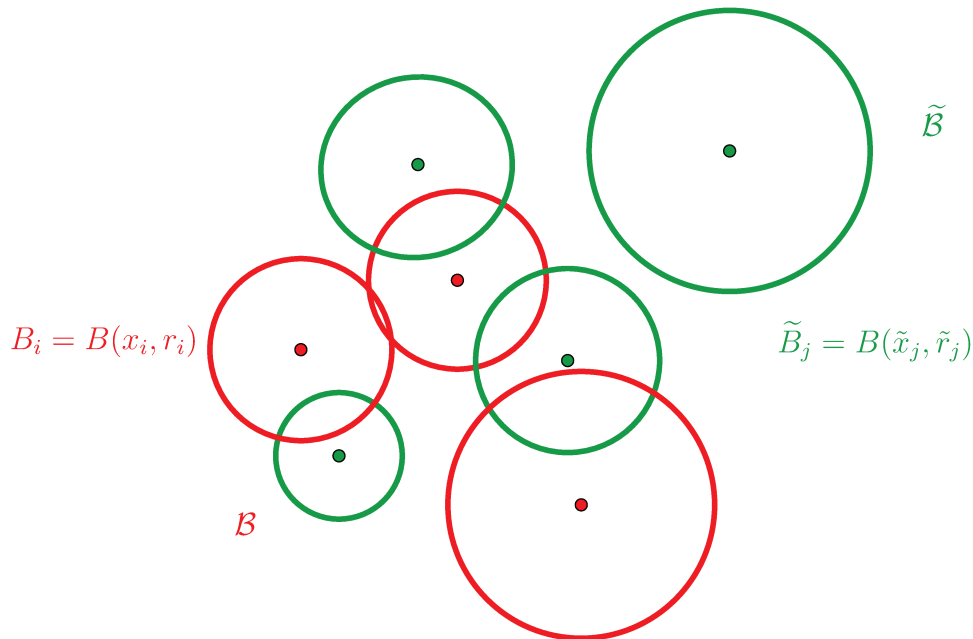


FIGURA 9.2.1. Cubrimiento por bolas de los conjuntos X e Y vistos como espacios métricos con atributos dados por los radios de las bolas.

Tomamos $X = \{x_i : i = 1, \dots, n\}$, $Y = \{\tilde{x}_j : j = 1, \dots, m\}$, d_X y d_Y las heredadas de \mathbb{R}^2 por los conjuntos X e Y respectivamente. Sean $\mu = \sum_{i=1}^n r_i \delta_{x_i}$ y $\nu = \sum_{j=1}^m \tilde{r}_j \delta_{\tilde{x}_j}$ las medidas probabilísticas que describen el atributo “radio” en cada vértice de cada grafo. Entonces (X, d_X, μ) e (Y, d_Y, ν) son espacios métricos de probabilidad. Por consiguiente $\mathfrak{d}_{GH}^K((X, d_X, \mu), (Y, d_Y, \nu))$ está bien definida. Para obtener una estimación superior de esta distancia podemos usar las inmersiones isométricas que definen las inclusiones de X e Y en \mathbb{R}^2 . Es decir

$$\begin{aligned} f : X &\rightarrow \mathbb{R}^2, \quad f(x_i) = x_i; \\ g : Y &\rightarrow \mathbb{R}^2, \quad g(\tilde{x}_j) = \tilde{x}_j. \end{aligned}$$

Entonces

$$\mathfrak{d}_{GH}^{KR}((X, d_X, \mu), (Y, d_Y, \nu)) \leq d_{H\mathbb{R}^2}(X, Y) + d_{K\mathbb{R}^2}(\mu, \nu),$$

donde $d_{H\mathbb{R}^2}$ denota la distancia de Hausdorff usual en el plano y $d_{K\mathbb{R}^2}$ la distancia de Kantorovich-Wasserstein entre probabilidades en \mathbb{R}^2 .

El mismo esquema puede repetirse con cualquier otra de las métricas entre medidas introducida en el Capítulo 7.

Un caso particular de la situación general en la que hemos definido la métrica \mathfrak{d}_{GH}^{KR} ha sido considerado en [ACI09].

9.3. Aplicación del enfoque de Hausdorff-Gromov a SUBE en AMBA

En la Sección 8.2 como aplicación de las métricas de tipo Gromov-Lipschitz estimamos las distancias entre los modelos de SUBE en AMBA con los distintos atributos normalizados en los vértices (distritos de AMBA) como medidas probabilísticas. La ventaja del enfoque de esta sección es que la métrica de Gromov-Hausdorff no requiere la misma cardinalidad de los dos espacios. Por lo tanto esto nos permite retomar, ahora haciendo entrar en juego los atributos de los distritos, en el ejemplo de la Sección 6.6 del Capítulo 6. más precisamente, ordenabamos los 41 distritos del AMBA por población decreciente y obtuvimos los distritos indexados de 101 a 141 en la pagina 96. Luego dado $r > 0$ construimos un cubrimiento disjunto del conjunto de los 41 distritos de la siguiente

forma:

$$A_{101}(r) = B_{d_1}(101, r)$$

$$A_{102}(r) = B_{d_1}(102, r) \setminus B_{d_1}(101, r)$$

$$A_{103}(r) = B_{d_1}(103, r) \setminus (B_{d_1}(101, r) \cup B_{d_1}(102, r))$$

⋮

$$A_j(r) = B_{d_1}(j, r) \setminus \left(\bigcup_{l=101}^{j-1} B_{d_1}(l, r) \right); \quad j \leq 141.$$

El algoritmo toma la forma

$$A_{101}(r) = \{101 \leq l \leq 141 : d_1(101, l) < r\}$$

$$A_{102}(r) = \{101 \leq l \leq 141 : d_1(102, l) < r \wedge d_1(101, l) \geq r\}$$

⋮

$$A_j(r) = \{101 \leq l \leq 141 : d_1(j, l) < r \wedge d_1(m, l) \geq r \forall m : 101 \leq m \leq j - 1\}; \quad j \leq 141.$$

Luego se eliminan los que son vacíos y se obtiene un cubrimiento disjunto de $\{1, \dots, 41\}$ dado por $\{A_j(r) : j \in I(r)\}$ donde $I(r) \subseteq \{1, \dots, 41\}$. Es decir $\{1, \dots, 41\} = \bigsqcup_{j \in I(r)} A_j(r)$. Para cada r entre 0 y 1 y para cada $j \in I(r)$ elegimos la ciudad con mayor población dentro de $A_j(r)$. Es decir tomamos en $A_j(r)$ un elemento $k(j, r) \in I(r)$ tal que el número de habitantes de la ciudad $k(j, r)$ es mayor que el de todas las otras ciudades que conforman el clúster $A_j(r)$. Con esto construimos un nuevo grafo $\mathcal{G}_2(r)$ para cada r dado por $\mathcal{G}_2(r) = (\mathcal{V}_2(r), d_1)$ con $\mathcal{V}_2(r) = \{k(j, r) : j \in I(r)\}$ y d_1 como antes, es la métrica difusiva para $t = 100$. Por otra parte los atributos de cada $K(j, r)$ para $j \in I(r)$ deben reunir todos los atributos de los distritos originales que $K(j, r)$ representa. Es decir, para $j \in I(r)$, $\tilde{\mu}_i(\{j\}) = \sum_{l \in I(r)} \mu_i(\{l\})$. Con esto obtenemos cuatro nuevos grafos para cada $r > 0$: $(\mathcal{V}_2(r), d_1, \tilde{\mu}_i)$ $i = 1, 2, 3, 4$ Usando Python obtenemos las siguientes curvas como estimaciones superiores para $d_{GH}^K((\mathcal{V}_2(r), d_1, \tilde{\mu}_i), (\mathcal{V}_1, d_1, \nu_i))$.

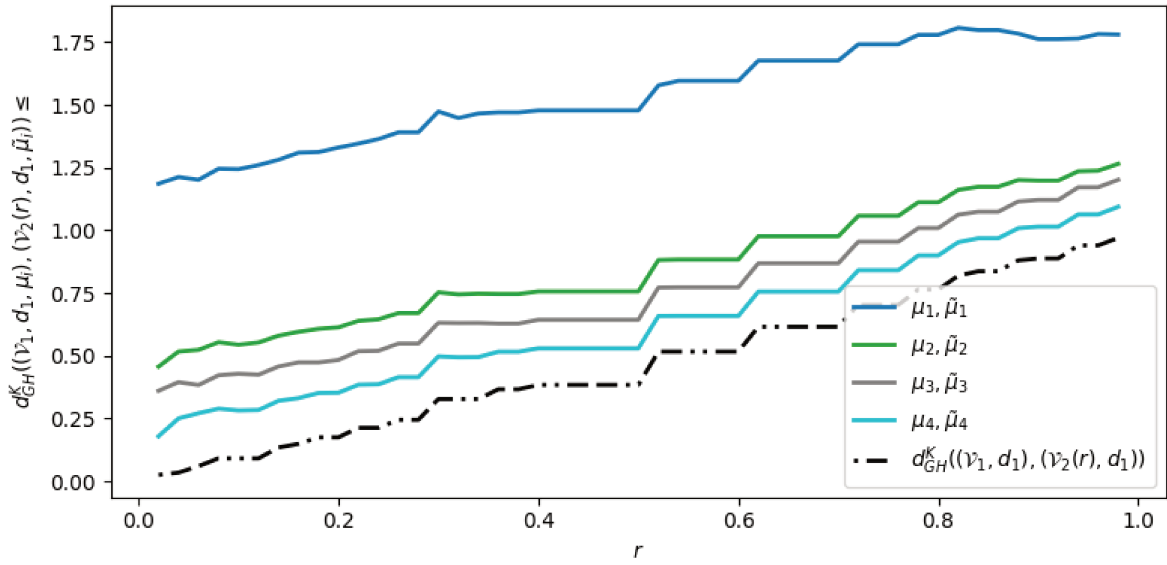


FIGURA 9.3.2. Estimaciones de $d_{GH}^{KR}((\mathcal{V}_1, d_1, \mu_i), (\mathcal{V}_2(r), d_1, \tilde{\mu}_i))$ $i = 1, 2, 3, 4$.

Observamos que, de nuevo, la distribución uniforme, de color azul muestra la mayor diferencia. Notamos también que la curva más baja de color cian (celeste intenso) es la de distancia entre $(\mathcal{V}_1, d_1, \mu_4)$ y $(\mathcal{V}_2(r), d_1, \tilde{\mu}_4)$, y que justamente μ_4 , y por lo tanto $\tilde{\mu}_4$, son las que registran la densidad de infectados por COVID-19 durante Julio de 2020.

Distancia entre grafos de afinidades con atributos. Estabilidad del espectro del Laplaciano

10.1. Introducción

En los capítulos anteriores hemos introducido e ilustrado, con ejemplos y algunas aplicaciones, varias nociones de distancias entre espacios métricos con medidas probabilísticas. En el Capítulo 2 vimos que bajo algunas hipótesis adicionales de transitividad en una afinidad, se tiene una métrica que la determina. Pero también vimos en nuestra aplicación al análisis del grafo de transporte público en AMBA, que la propiedad de transitividad está lejos de verificarse en situaciones reales. Esto hace que las nociones de distancias entre grafos métricos no estén siempre a disposición cuando tenemos grafos de afinidades. Será conveniente entonces, introducir también alguna noción útil y simple de métrica entre grafos de afinidades que tengan atributos probabilísticos en los vértices. Algunas de las razones para hacerlo son el análisis de estabilidad del espectro del Laplaciano y su uso en el próximo capítulo para introducir variables aleatorias valuadas en grafos de afinidades con atributos.

10.2. Distancia entre grafos con afinidades y atributos probabilísticos

El Teorema de Estabilidad del Espectro de matrices simétricas (ver [Wil88] Capítulo 2) sugiere que una métrica útil entre las afinidades de un grafo puede darse en términos de la norma de Frobenius de la diferencia de las matrices de afinidad. Por otra parte, siendo los atributos medidas de probabilidad en los vértices del grafo cualquier métrica entre medidas puede usarse adecuadamente para construir distancias entre espacios con afinidad y atributos. Nos limitaremos a Kantorovich-Rubinstein y a la métrica euclídea natural entre atributos.

Sea $\mathcal{V} = \{1, 2, \dots, n\}$, $n \in \mathbb{N}$. Sea $\mathcal{G}_1 = (\mathcal{V}, \mathcal{E}, \bar{a}^1, W^1)$ y $\mathcal{G}_2 = (\mathcal{V}, \mathcal{E}, \bar{a}^2, W^2)$, con $\bar{a}^1, \bar{a}^2 \in S = \{\bar{a} = (a_1, \dots, a_n) \in \mathbb{R}^n : a_i > 0 \text{ y } \sum_{i=1}^n a_i = 1\}$ dos grafos ponderados sobre el mismo conjunto \mathcal{V} . Si en \mathcal{V} hay una métrica dada ρ , entonces está bien definida la distancia

de Kantorovich-Rubinstein entre medidas de probabilidad y podemos definir

$$d_F^{KR}(\mathcal{G}_1, \mathcal{G}_2) = \|W^1 - W^2\|_F + d_{KR}(\mu_{\bar{a}^1}, \mu_{\bar{a}^2}),$$

donde $\mu_{\bar{a}^i}(E) = \sum_{j \in E} a_j^i$ con $i = 1, 2$. Explícitamente,

$$d_F^{KR}(\mathcal{G}_1, \mathcal{G}_2) = \left(\sum_{i,j=1}^n (w_{ij}^1 - w_{ij}^2)^2 \right)^{\frac{1}{2}} + \sup \left\{ \left| \int_{\mathcal{V}} \psi d\mu_{\bar{a}^1} - \int_{\mathcal{V}} \psi d\mu_{\bar{a}^2} \right| : \frac{|\psi(i) - \psi(j)|}{\rho(i, j)} \leq 1 \right\}.$$

Cuando en lugar de la métrica de Kantorovich entre $\mu_{\bar{a}^1}$ y $\mu_{\bar{a}^2}$ usamos la distancia Euclídea entre \bar{a}^1 y \bar{a}^2 tenemos

$$d_F(\mathcal{G}_1, \mathcal{G}_2) = \|W^1 - W^2\|_F + \|\bar{a}^1 - \bar{a}^2\|,$$

donde $\|\bar{a}^1 - \bar{a}^2\|^2 = \sum_{i=1}^n (a_i^1 - a_i^2)^2$. Notemos que en esta métrica euclídea de los vectores \bar{a} no estamos usando que la suma de los a_i sea uno.

Por el Corolario 7.7.2, tenemos que $d_F^{KR} \leq \sqrt{n} \text{diam}_\rho \mathcal{V} d_F$.

10.3. Estabilidad del espectro del Laplaciano

La siguiente desigualdad es un resultado clásico del álgebra lineal, y su demostración puede verse en [Wil88].

TEOREMA 10.3.1. *Sea M una matriz real y simétrica. Entonces existe un orden de autovalores $\{\lambda_i\}$ de M y $\{\tilde{\lambda}_i\}$ de $M + E$, tal que*

$$\sum_{i=1}^n |\lambda_i - \tilde{\lambda}_i|^2 \leq \|E\|_F^2,$$

donde la norma $\|E\|_F^2 = \sum_{j=1}^n \sum_{i=1}^n |E_{ji}|^2$ es la norma de Frobenius de la matriz E_{ij} .

A partir de este teorema obtendremos un resultado para la aproximación del espectro del Laplaciano de un grafo ponderado. Usaremos aquí la versión del Laplaciano asociado al grafo $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \bar{a}, W)$ que introdujimos en la Sección 1.3 del Capítulo 1,

$$(\Delta_{\mathcal{G}} f)_i = \frac{1}{a_i} \sum_{j=1}^n w_{ij} (f_j - f_i),$$

que en forma matricial se escribe

$$\Delta_{\mathcal{G}} = A^{-1}(W - D),$$

donde A es la matriz diagonal tal que $A_{ii} = a_i > 0$, con $i = 1, \dots, n$ y D es la matriz diagonal $D_{ii} = \sum_{j=1}^n w_{ij}$, con $i = 1, \dots, n$. Los resultados del Capítulo 1, Sección 1.3, muestran que $\Delta_{\mathcal{G}}$ es una matriz real simétrica semidefinida negativa. Por consiguiente podemos aplicar el Teorema 10.3.1.

TEOREMA 10.3.2. *Sea K un compacto en $(\mathbb{R}_+)^n = \mathbb{R}_+ \times \dots \times \mathbb{R}_+$. Entonces existe una constante $C_{K,n}$ que depende de K y n tal que para $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \bar{a}, W)$ y $\tilde{\mathcal{G}} = (\mathcal{V}, \mathcal{E}, \tilde{a}, \tilde{W})$ con \bar{a} y $\tilde{a} \in K$, $w_{ij} \geq 0$, $\tilde{w}_{ij} \geq 0$, se tiene que los autovalores de los Laplacianos $\Delta_{\mathcal{G}}$ y $\Delta_{\tilde{\mathcal{G}}}$, adecuadamente ordenados, satisfacen la siguiente estimación de tipo Lipschitz con respecto a afinidades y atributos como variables,*

$$\left(\sum_{i=1}^n |\lambda_i - \tilde{\lambda}_i|^2 \right)^{\frac{1}{2}} \leq C_{K,n} \left(\|\bar{a} - \tilde{a}\|_2 \|W - D\|_F + \|\tilde{W} - W\|_F \right),$$

donde $\|\cdot\|_2$ denota la norma euclídea del vector.

DEMOSTRACIÓN. Notemos en primer lugar que

(10.1)

$$\begin{aligned} \|\Delta_{\mathcal{G}} - \Delta_{\tilde{\mathcal{G}}}\|_F &= \|A^{-1}(W - D) - \tilde{A}^{-1}(\tilde{W} - \tilde{D})\|_F \\ &= \|A^{-1}(W - D) - \tilde{A}^{-1}(W - D) + \tilde{A}^{-1}(W - D) - \tilde{A}^{-1}(\tilde{W} - \tilde{D})\|_F \\ &= \left\| (A^{-1} - \tilde{A}^{-1})(W - D) + \tilde{A}^{-1} \left[(W - D) - (\tilde{W} - \tilde{D}) \right] \right\|_F \\ &= \left\| (A^{-1} - \tilde{A}^{-1})(W - D) + \tilde{A}^{-1} \left[(\tilde{D} - D) + (W - \tilde{W}) \right] \right\|_F \\ &\leq \|(A^{-1} - \tilde{A}^{-1})(W - D)\|_F + \|\tilde{A}^{-1}\|_F \left(\|\tilde{D} - D\|_F + \|W - \tilde{W}\|_F \right). \end{aligned}$$

Estudiemos el lado derecho de esta desigualdad tratando los dos términos por separado. Primero,

$$\begin{aligned} (10.2) \quad \|(A^{-1} - \tilde{A}^{-1})(W - D)\|_F &\leq \|(A^{-1} - \tilde{A}^{-1})\|_F \|W - D\|_F \\ &= \left(\sum_{i=1}^n \left| \frac{1}{a_i} - \frac{1}{\tilde{a}_i} \right|^2 \right)^{\frac{1}{2}} \|W - D\|_F \\ &= \left(\sum_{i=1}^n \left(\frac{|\tilde{a}_i - a_i|}{a_i \tilde{a}_i} \right)^2 \right)^{\frac{1}{2}} \|W - D\|_F \end{aligned}$$

$$\begin{aligned} &\leq \frac{1}{\min |a_i| \min |\tilde{a}_i|} \|\tilde{\mathbf{a}} - \mathbf{a}\|_2 \|W - D\|_F \\ &= C_K \|\tilde{\mathbf{a}} - \mathbf{a}\|_2 \|W - D\|_F, \end{aligned}$$

donde C_K es una constante que depende del conjunto compacto K en $(\mathbb{R}_+)^n$. Para el segundo termino en el lado derecho de (10.1),

$$(10.3) \quad \|\tilde{A}^{-1}\|_F \left(\|\tilde{D} - D\|_F + \|W - \tilde{W}\|_F \right) = \|\tilde{A}^{-1}\|_F \|\tilde{D} - D\|_F + \|\tilde{A}^{-1}\|_F \|W - \tilde{W}\|_F.$$

Estimemos el primer sumando a la derecha en (10.3),

$$\|\tilde{A}^{-1}\|_F \|\tilde{D} - D\|_F \leq \|\tilde{A}^{-1}\|_F \left(\sum_{i=1}^n \left[\sum_{j=1}^n (w_{ij} - \tilde{w}_{ij}) \right]^2 \right)^{\frac{1}{2}}.$$

Por la desigualdad de Schwartz

$$\begin{aligned} \|\tilde{A}^{-1}\|_F \|\tilde{D} - D\| &\leq \|\tilde{A}^{-1}\|_F \left(\sum_{i=1}^n \left[\left(\sum_{j=1}^n 1^2 \right)^{\frac{1}{2}} \left(\sum_{j=1}^n (w_{ij} - \tilde{w}_{ij})^2 \right)^{\frac{1}{2}} \right]^2 \right)^{\frac{1}{2}} \\ &= \|\tilde{A}^{-1}\|_F \left(\sum_{i=1}^n \left[\sqrt{n} \left(\sum_{j=1}^n (w_{ij} - \tilde{w}_{ij})^2 \right)^{\frac{1}{2}} \right]^2 \right)^{\frac{1}{2}} \\ &= \|\tilde{A}^{-1}\|_F \left(\sum_{i=1}^n n \sum_{j=1}^n (w_{ij} - \tilde{w}_{ij})^2 \right)^{\frac{1}{2}} \\ &\leq \|\tilde{A}^{-1}\|_F \sqrt{n} \|W - \tilde{W}\|_F. \end{aligned}$$

Entonces, volviendo a (10.3) con esta estimación,

$$(10.4) \quad \|\tilde{A}^{-1}\|_F \left(\|\tilde{D} - D\|_F + \|W - \tilde{W}\|_F \right) \leq \|\tilde{A}^{-1}\|_F \left(\sqrt{n} \|W - \tilde{W}\|_F + \|W - \tilde{W}\|_F \right).$$

Con las estimaciones (10.2) y (10.4) en el miembro derecho de (10.1), tenemos finalmente

$$\|\Delta_{\mathcal{G}} - \Delta_{\tilde{\mathcal{G}}}\|_F \leq C_{K,n} \left(\|\mathbf{a} - \tilde{\mathbf{a}}\|_2 \|W - D\|_F + \|\tilde{W} - W\|_F \right),$$

usando el Teorema 10.3.1 obtenemos el resultado, es decir

$$\left(\sum_{i=1}^n |\lambda_i - \tilde{\lambda}_i|^2 \right)^{\frac{1}{2}} \leq C_{K,n} \left(\|\bar{a} - \tilde{a}\|_2 \|W - D\|_F + \|\widetilde{W} - W\|_F \right),$$

para alguna ordenación de los autovalores. \square

Observamos que la constante de la desigualdad del Teorema 10.3.2 depende de la distancia de K a los planos coordenados de $(\mathbb{R}_+)^n$ y es cada vez más grande cuando esta distancia es más pequeña. Cuando aplicamos este resultado a nuestro caso particular en que $\bar{a} \in S$ será importante que ninguno de los vértices en \mathcal{V} tenga peso nulo.

10.4. Un ejemplo, perturbaciones del grafo bipartito

A continuación ilustramos, a través de perturbaciones del grafo bipartito de la Sección 1.4 del Capítulo 1, cómo se modifica la secuencia de autovalores perturbando \bar{a} , W , y ambos a la vez. De las cuatro gráficas contenidas en la Figura 10.4.1, la primera contiene los autovalores obtenidos en la Sección 1.4 para el grafo bipartito $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$, $\mathcal{V}_1 = \{1, 2, 3\}$ y $\mathcal{V}_2 = \{4, 5, 6, 7, 8, 9\}$. La segunda gráfica se obtiene perturbando sólo el vector \bar{a} de manera aleatoria (uniforme, usando la función `random.uniform` de Python) en el intervalo $(\frac{9}{10}, \frac{11}{10})$ y luego normalizando a probabilidad para obtener vectores \tilde{a} . La tercera gráfica se obtiene perturbando sólo la matriz W a una nueva \widetilde{W} obtenida de manera aleatoria y uniforme eligiendo los \tilde{w}_{ij} en el intervalo $(\frac{9}{10}, \frac{11}{10})$ o en el intervalo $(0, \frac{1}{10})$, según sea $w_{ij} = 1$ o $w_{ij} = 0$, y luego normalizando a probabilidad. La gráfica 4 muestra el efecto de la ejecución simultánea de ambas perturbaciones \tilde{a} y \widetilde{W} de \bar{a} y de W respectivamente.

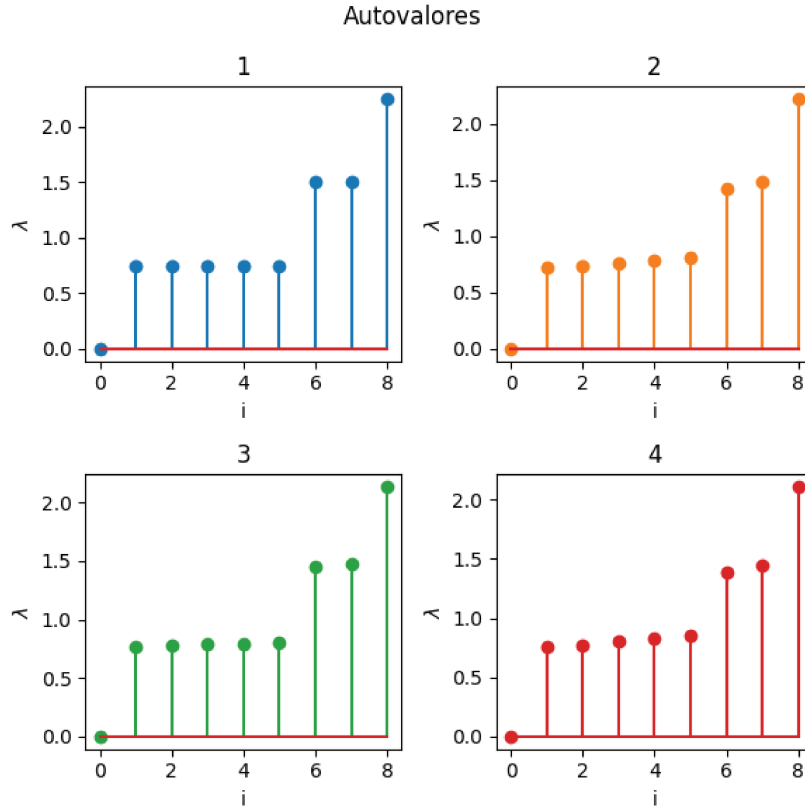


FIGURA 10.4.1. Perturbaciones de afinidades y atributos de un grafo bipartito y perturbación del espectro.

Para $j = 1, 2, 3, 4$, denotamos $\{\lambda_i^j : i = 1, \dots, 9\}$ a los autovalores correspondientes a cada una de las cuatro gráficas en la Figura 10.4.1. Calculando las diferencias de los autovalores de los casos 2, 3 y 4 con los del grafo original 1, tenemos

$$\text{a) } \sum_{i=1}^n (\lambda_i^1 - \lambda_i^2)^2 \leq 0.0129$$

$$\text{b) } \sum_{i=1}^n (\lambda_i^1 - \lambda_i^3)^2 \leq 0.0248$$

$$\text{c) } \sum_{i=1}^n (\lambda_i^1 - \lambda_i^4)^2 \leq 0.0565,$$

Por otra parte, por construcción de las aproximaciones tenemos

$$\text{a') } \|\bar{a}^1 - \bar{a}^2\|_2 \leq 0.019 \quad \text{y} \quad \|W^1 - W^2\|_F = 0$$

$$\text{b') } \|\bar{a}^1 - \bar{a}^3\|_2 = 0 \quad \text{y} \quad \|W^1 - W^3\|_F \leq 0.013$$

$$\text{c') } \|\bar{a}^1 - \bar{a}^4\|_2 \leq 0.019 \quad \text{y} \quad \|W^1 - W^4\|_F \leq 0.013,$$

esto es consistente con la desigualdad obtenida en el Teorema 10.3.2.

10.5. Aplicación al cálculo de las distancias entre las afinidades dadas por el transporte público en AMBA antes y durante ASPO-2020 por COVID-19

Entre los datos que el grupo LABRA obtuvo de Nación Servicios para el sistema SUBE de transporte público en AMBA, disponemos de matrices de conectividad para varios días distintos de aquel inicio de la pandemia COVID-19 en Argentina. Aquí sólo mostramos tres de esos días.



El 4 de marzo, prepandémico, el 8 de abril, a menos de 20 días del inicio del Aislamiento Social Preventivo y Obligatorio por Decreto Presidencial 297/2020, y, finalmente, el 3 de junio de 2020 un par de días antes de la extensión de la cuarentena y después de la sanción del Decreto 459/2020 que facultaba a gobiernos locales a debilitar las condiciones de aislamiento.

Las tres matrices son las siguientes

Matrix SUBE 4 de Marzo 2020, W_M. A large grid of numerical data representing connectivity between various points in the AMBA system.

FIGURA 10.5.2. Matriz SUBE 4 de Marzo 2020, W_M.

| Distancia | W_M, W_A | W_M, W_J | W_A, W_J |
|------------|------------|------------|------------|
| d_F | 203836 | 191360 | 13300 |
| d_1 | 2083474 | 1951206 | 135456 |
| d_∞ | 77439 | 72011 | 5428 |

CUADRO 1. Distancias entre afinidades de transporte en distintos momentos de 2020.

donde W es alguna de esas tres matrices W_M , W_A o W_J y \bar{a} es alguno de esos vectores de atributos. En estas situaciones podemos calcular d_F entre dos cualquiera de esos grafos.

Los casos más interesantes desde el punto de vista epidemiológico, que es el que dio origen a estas consideraciones, son aquellos en los que los pesos a_i en los 41 distritos reflejan de algún modo el estado de evolución de COVID-19 en cada uno de ellos para las épocas en las que disponemos de los datos de transporte público a través de SUBE en AMBA. Para los casos expuestos arriba de las tres matrices W_M , W_A y W_J , sólo las dos últimas son posteriores al inicio de la pandemia en la Argentina. Con la ayuda del Dr. Franco Marsico de la UBA, obtuvimos del Ministerio de Salud de la Nación, los datos acumulados de infectados en cada uno de los 41 distritos para los meses de Abril y de Junio de 2020. Los datos obtenidos se pueden ver en el Cuadro 2. La Segunda columna contiene las componentes del vector \bar{a}_A^I y la tercera las del vector \bar{a}_J^I .

Resulta sencillo estimar la distancia d_F entre los grafos $\mathcal{G}_A^I = (\mathcal{V}, \mathcal{E}, \bar{a}_A^I, W_A)$ y $\mathcal{G}_J^I = (\mathcal{V}, \mathcal{E}, \bar{a}_J^I, W_J)$. Si en vez de atributos asociados a niveles de infectados por distrito, usamos atributos uniformes, la distancia d_F entre los correspondientes grafos $\mathcal{G}_A^U = (\mathcal{V}, \mathcal{E}, \bar{a}_A^U, W_A)$ y $\mathcal{G}_J^U = (\mathcal{V}, \mathcal{E}, \bar{a}_J^U, W_J)$ está dada simplemente por el valor de $d_F(W_A, W_J) = 13300$ que figura en el Cuadro 1. Para los atributos de infectados por distritos, tenemos

$$d_F = \|W_A - W_J\|_F + \|\bar{a}_A^I - \bar{a}_J^I\|_2 \simeq 33539.$$

En el espectro de los Laplacianos correspondientes a \mathcal{G}_A^U , \mathcal{G}_J^U , \mathcal{G}_A^I y \mathcal{G}_J^I se pone de manifiesto la relevancia de cada una de las ponderaciones: aristas por SUBE y vértices por cantidad de infectados.

| Distrito | Abril | Junio |
|--------------------|-------|-------|
| Almirante Brown | 42 | 1411 |
| Berisso | 3 | 164 |
| Campana | 5 | 101 |
| Escobar | 39 | 406 |
| Ezeiza | 35 | 353 |
| General Rodríguez | 6 | 142 |
| Ituzaingó | 21 | 266 |
| Lanús | 31 | 1601 |
| Luján | 15 | 121 |
| Moreno | 24 | 807 |
| Pilar | 61 | 513 |
| San Fernando | 24 | 297 |
| San Vicente | 25 | 99 |
| Vicente López | 47 | 377 |
| Avellaneda | 48 | 1775 |
| Brandsen | 7 | 35 |
| Cañuelas | 1 | 83 |
| Esteban Echevarría | 17 | 694 |
| Florencio Varela | 35 | 892 |
| General San Martín | 111 | 1253 |

| Distrito | Abril | Junio |
|-----------------------|-------|-------|
| José C. Paz | 17 | 452 |
| La Plata | 45 | 671 |
| Marcos Paz | 2 | 95 |
| Merlo | 75 | 958 |
| Presidente Perón | 3 | 191 |
| San Isidro | 29 | 591 |
| Tigre | 49 | 654 |
| Zárate | 5 | 129 |
| Berazategui | 24 | 726 |
| CABA | 862 | 20068 |
| Ensenada | 3 | 95 |
| Exaltación de la Cruz | 1 | 21 |
| General Las Heras | 1 | 50 |
| Hurlingham | 16 | 420 |
| La Matanza | 112 | 4371 |
| Lomas de Zamora | 37 | 1767 |
| Malvinas Argentinas | 33 | 415 |
| Morón | 44 | 594 |
| Quilmes | 69 | 2110 |
| San Miguel | 28 | 590 |
| Tres de Febrero | 40 | 1123 |

CUADRO 2. Infectados por distrito en Abril y Junio de 2020.

Observamos en la Figura 10.5.5 que, aunque las dos gráficas no están en las mismas escalas, se advierte el predominio espectral del transporte por sobre el número de infectados por distrito al principio de la pandemia (Abril), mientras que luego, en Junio el impacto del número de infectados por distrito es mayor y los autovalores se invierten.

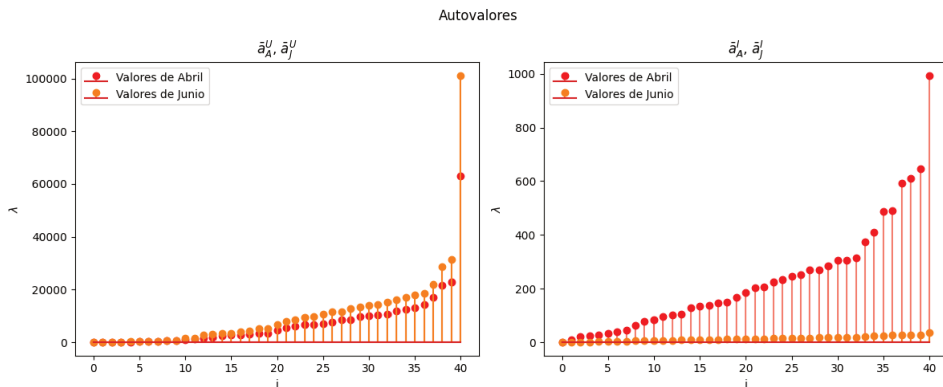


FIGURA 10.5.5. Comparaciones del espectro entre atributos uniformes y número de infectados en Abril y Junio de 2020.

Así por ejemplo si calculamos las distancias difusivas $d_{A,t}^I$ y $d_{J,t}^I$ entre Ezeiza (=4) y CABA (=29) para $t = 0.01$ tenemos

$$d_{A,t}^I(4, 29) = 0.4187225 \quad y \quad d_{J,t}^I(4, 29) = 0.96957266.$$

VARIABLES ALEATORIAS VALUADAS EN ESPACIOS MÉTRICOS CON MEDIDAS

11.1. Introducción

En el Capítulo 8 hemos dotado a $\mathcal{M}^l(X_0, d_0)$ con una estructura de espacio casi-métrico primero y métrico después y esto nos permite considerar objetos aleatorios en esta super-estructura. Es decir, variables aleatorias valuadas en espacios métricos probabilísticos. Formalizamos esta idea en la Sección 11.2 y luego iremos al caso particular, pero importante, de los grafos métricos con atributos en los que tanto las métricas como los atributos son variables aleatorias. También consideramos variables aleatorias valuadas en grafos con afinidades y atributos. Finalmente ilustraremos la aleatoriedad en el espectro de grafos clásicos y del grafo de nuestra aplicación a SUBE en AMBA.

11.2. Variables aleatorias valuadas en $(\mathcal{M}^l(X_0, d_0), \mathfrak{d}_{GL}^{KR})$

La topología generada por la casi-métrica \mathfrak{d}_{GL}^{KR} del Capítulo 8 nos permite definir Borelianos en $\mathcal{M}^l(X_0, d_0)$ y entonces variables aleatorias.

DEFINICIÓN 11.2.1. Sea $(\Omega, \mathcal{F}, \mathcal{P})$ un espacio de probabilidad. Una función $X : \Omega \rightarrow \mathcal{M}^l(X_0, d_0)$ es una **variable aleatoria con valores en $\mathcal{M}^l(X_0, d_0)$** si

$$X^{-1}(\mathcal{A}) = \{\omega \in \Omega : X(\omega) \in \mathcal{A}\}$$

es un elemento de \mathcal{F} para cada \mathcal{A} de Borel en $\mathcal{M}^l(X_0, d_0)$. La **distribución de X** es la **medida** de Borel ν_X en $\mathcal{M}^l(X_0, d_0)$ definida por la probabilidad de la preimagen por X de \mathcal{A} , precisamente

$$\nu_X(\mathcal{A}) = \mathcal{P}(X^{-1}(\mathcal{A})),$$

con \mathcal{A} Boreliano en $\mathcal{M}^l(X_0, d_0)$.

Dado un elemento $\mathbb{M}_1 = (X_1, d_1, \mu_1) \in \mathcal{M}^l(X_0, d_0)$, y $m \geq 1$, la función $\varphi_m : \mathcal{M}^l(X_0, d_0) \rightarrow \mathbb{R}^+ \cup \{0\}$ dada por $\varphi_m(\mathbb{M}) = [\mathfrak{d}_{GL}^{KR}(\mathbb{M}_1, \mathbb{M})]^m$ es continua y no negativa. Por consiguiente están bien definidos, posiblemente infinitos, los momentos de orden m

de una variable aleatoria $X : \Omega \rightarrow \mathcal{M}^l(X_0, d_0)$, alrededor de \mathbb{M}_1 , por

$$\sigma_m^{\mathbb{M}_1}(X) = \int_{\Omega} [\mathfrak{d}_{GL}^{KR}(\mathbb{M}_1, X(\omega))]^m d\mathcal{P}(\omega).$$

PROPOSICIÓN 11.2.2. Sean $X : \Omega \rightarrow \mathcal{M}^l(X_0, d_0)$ una variable aleatoria en $\mathcal{M}^l(X_0, d_0)$ y m entero positivo, entonces

$$\sigma_l^{\mathbb{M}_1}(X) = \int_{\mathbb{M} \in \mathcal{M}^l(X_0, d_0)} [\mathfrak{d}_{GL}^K(\mathbb{M}_1, \mathbb{M})]^l d\nu_X(\mathbb{M}).$$

DEMOSTRACIÓN. Sea \mathcal{A} un boreliano en $\mathcal{M}^l(X_0, d_0)$ y sea $\varphi : \mathcal{M}^l(X_0, d_0) \rightarrow \mathbb{R}^+ \cup \{0\}$ dada por $\varphi(\mathbb{M}) = \chi_{\mathcal{A}}(\mathbb{M})$ que, como es usual, vale 1 si $\mathbb{M} \in \mathcal{A}$ y 0 en caso contrario. Entonces

$$\begin{aligned} \int_{\Omega} \varphi(X(\omega)) d\mathcal{P}(\omega) &= \int_{\Omega} \chi_{\mathcal{A}}(X(\omega)) d\mathcal{P}(\omega) \\ &= \int_{\{\omega \in \Omega : X(\omega) \in \mathcal{A}\}} d\mathcal{P}(\omega) \\ &= \mathcal{P}(\{\omega \in \Omega : X(\omega) \in \mathcal{A}\}) \\ &= \nu_X(\mathcal{A}) \\ &= \int_{\mathbb{M} \in \mathcal{A}} d\nu_X(\mathbb{M}) \\ &= \int_{\mathcal{M}^l(X_0, d_0)} \chi_{\mathcal{A}}(\mathbb{M}) d\nu_X(\mathbb{M}) \\ &= \int_{\mathcal{M}^l(X_0, d_0)} \varphi(\mathbb{M}) d\nu_X(\mathbb{M}). \end{aligned}$$

Argumentos usuales de aproximación creciente de la función $[\mathfrak{d}_{GL}^{KR}(\mathbb{M}_1, \cdot)]^m$ por funciones simples prueban la proposición. \square

11.3. Grafos métricos aleatorios en un espacio métrico con atributos

Sea como antes (Ξ, ρ) un espacio métrico. Sea $\mathcal{V} = \{1, 2, \dots, n\}$, $n \in \mathbb{N}$. Sea $\{\xi_j^0 : j = 1, \dots, n\}$ un subconjunto $4R$ disperso de Ξ , es decir $\rho(\xi_j^0, \xi_k^0) \geq 4R$, para toda elección de $j \neq k$, entre 1 y n . Sea $\Xi^n = \Xi \times \Xi \times \dots \times \Xi$ el conjunto de las n -uplas $\bar{\xi} = (\xi_1, \dots, \xi_n)$ con $\xi_j \in \Xi$ sea $\rho^n(\bar{\xi}, \bar{\eta}) = \max\{\rho(\xi_i, \eta_i) : i = 1, \dots, n\}$. Sea $\bar{\xi}^0 \in \Xi^n$, dado por $\bar{\xi}^0 = (\xi_1^0, \dots, \xi_n^0)$

$$B_{\rho^n}(\bar{\xi}^0, R) = \{\bar{\xi} \in \Xi^n : \rho^n(\bar{\xi}, \bar{\xi}^0) < R\}$$

$$\begin{aligned}
 &= \{\bar{\xi} \in \Xi^n : \rho(\xi_i, \xi_i^0) < R \text{ para todo } i = 1, \dots, n\} \\
 &= \prod_{i=1}^n B_\rho(\xi_i^0, R).
 \end{aligned}$$

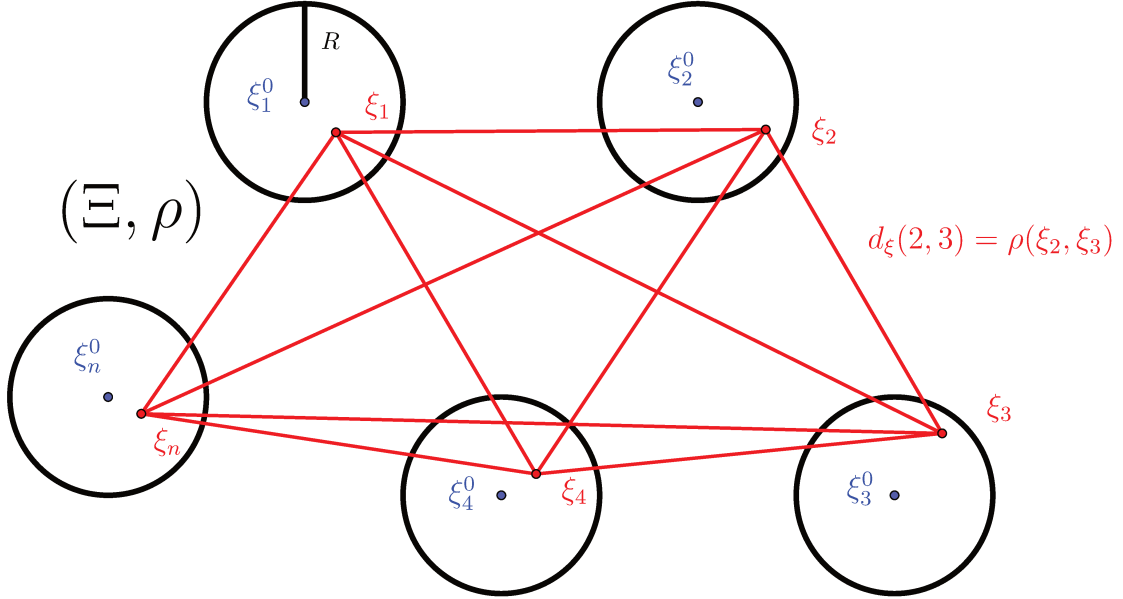


FIGURA 11.3.1. $(X_0, d_0) = (\mathcal{V}, d_0)$ y $(\mathcal{V}, d_{\bar{\xi}}) : \bar{\xi} \in B_{\rho^n}(\bar{\xi}^0, R)$.

Como antes tomamos una familia de métricas en $\mathcal{V} = \{1, \dots, n\}$ indexadas por $\bar{\xi} \in B_{\rho^n}(\bar{\xi}^0, R)$, de la siguiente manera. Si $\bar{\xi} \in B_{\rho^n}(\bar{\xi}^0, R)$ definimos

$$(11.1) \quad d_{\bar{\xi}}(i, j) = \rho(\xi_i, \xi_j), \quad i, j \in \mathcal{V}.$$

Consideramos ahora n -variables aleatorias $\xi_i : \Omega \rightarrow \Xi$, $i = 1, \dots, n$, valuadas en el espacio métrico (Ξ, ρ) tales que $\xi_i \in B_\rho(\xi_i^0, R)$. Denotamos con $\bar{\xi}(\omega)$ al vector aleatorio cuyas componentes son las $\xi_i(\omega)$. Sean también $Z_i : \Omega \rightarrow \mathbb{R}^+$, $i = 1, \dots, n$ n -variables aleatorias reales y positivas con $\sum_{i=1}^n Z_i(\omega) = 1$. Denotamos con $\bar{Z}(\omega)$ el vector aleatorio cuyas componentes son las $Z_i(\omega)$. Por (11.1), para cada $\omega \in \Omega$ se tiene $\delta(\omega)(i, j) = \rho(\xi_i(\omega), \xi_j(\omega))$, $i, j \in \mathcal{V}$ es una métrica en \mathcal{V} . Definimos en \mathcal{V} las medidas aleatorias dadas por $\mu(\omega)(E) = \sum_{i \in E} Z_i(\omega)$, $E \subseteq \mathcal{V}$, $\omega \in \Omega$.

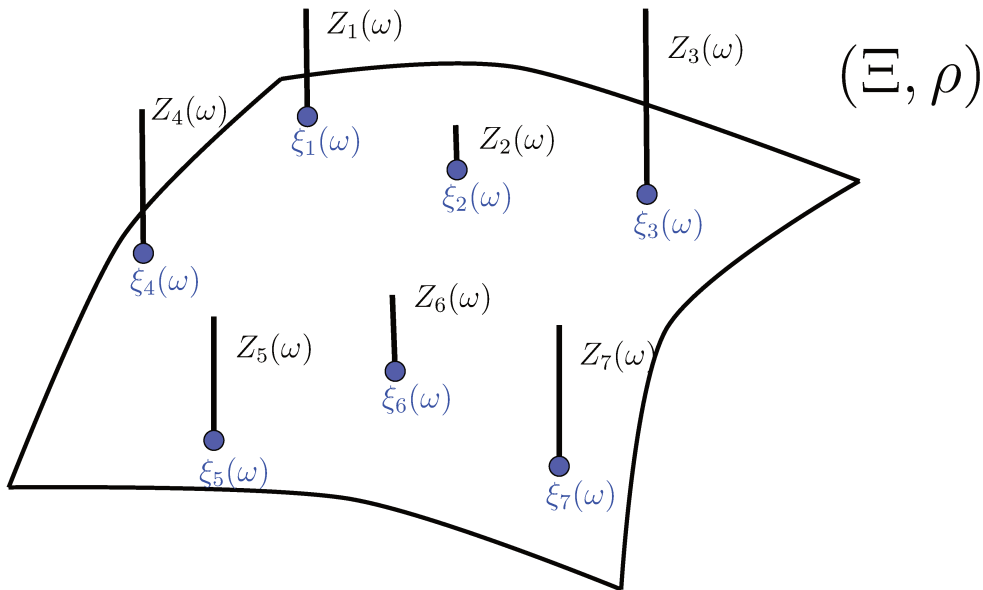


FIGURA 11.3.2. Puntos aleatorios en (Ξ, ρ) y atributos aleatorios en cada uno de ellos.

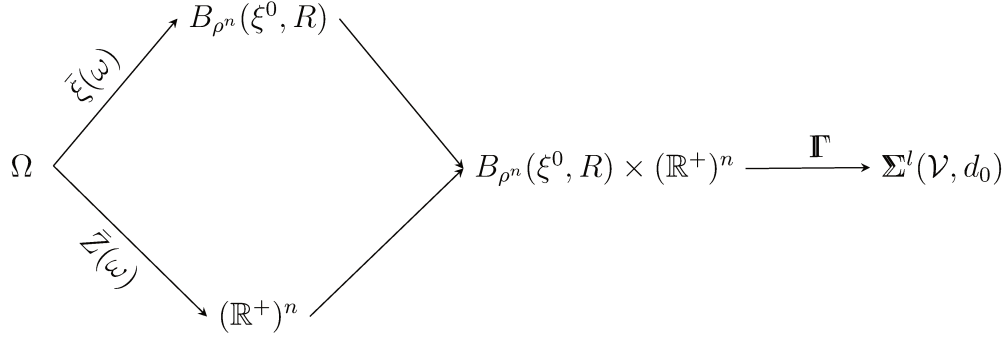
Estos dos objetos aleatorios $\bar{\xi}(\omega)$ y $\bar{Z}(\omega)$ nos permiten construir variables aleatorias valuadas en espacios métricos con medidas probabilísticas sobre el mismo conjunto \mathcal{V} por

$$\omega \in \Omega \mapsto X(\omega) = (\mathcal{V}, \delta(\omega), \mu(\omega)).$$

Notamos que X toma valores en un subconjunto particular de $\mathcal{M}^l(X_0, d_0)$ dado por $\Sigma^l(\mathcal{V}, d_0) = \{(\mathcal{V}, d_{\bar{\xi}}, \mu_{\bar{a}}) : \bar{\xi} \in B_{\rho^n}(\bar{\xi}^0, R) \text{ y } \bar{a} \in S\}$. Los resultados de continuidad introducidos paulatinamente en las Secciones 5.3, 7.7 y 8.3 nos permiten probar que, en este contexto, X es efectivamente una variable aleatoria.

PROPOSICIÓN 11.3.1. *Sea como en el Proposición 8.3.1 $l \geq \frac{D}{R}$. $X : \Omega \rightarrow \Sigma^l(\mathcal{V}, d_0)$ dada por $X(\omega) = (\mathcal{V}, \delta(\omega), \mu(\omega))$ es medible si en el codominio tomamos la σ -álgebra de Borel generada por la topología que d_{GL}^{KR} induce en $\Sigma^l(\mathcal{V}, d_0)$.*

DEMOSTRACIÓN. Notemos que



donde $\mathbf{\Gamma}(\bar{\xi}, \bar{a}) = (\mathcal{V}, d_{\bar{\xi}}, \mu_{\bar{a}})$ es continua por el Teorema 8.3.2. Entonces $X = \mathbf{\Gamma}(\bar{\xi}, \bar{Z})$. Puesto que si \mathcal{A} es un Boreliano en $\Sigma^l(\mathcal{V}, d_0)$ la continuidad de $\mathbf{\Gamma}$ implica que $\mathbf{\Gamma}^{-1}(\mathcal{A})$ también es Boreliano en $\Xi^n \times (\mathbb{R}^+)^n$, tenemos $X^{-1}(\mathcal{A}) = (\bar{\xi}, \bar{Z})^{-1}(\mathbf{\Gamma}^{-1}(\mathcal{A}))$ que está en la σ -álgebra de los eventos de Ω . \square

11.4. Variables aleatorias valuadas en grafos de afinidades con atributos

Sea $\mathcal{V} = \{1, 2, \dots, n\}$, con $n \in \mathbb{N}$, dado. Sea $(\Omega, \mathcal{F}, \mathcal{P})$ un espacio de probabilidad. Una función \mathcal{G} definida en Ω con valores en el conjunto de los grafos de afinidades con atributos en \mathcal{V} , que denotamos $\mathbb{G}(\mathcal{V})$, dada por $\mathcal{G}(\omega) = (\mathcal{V}, \mathcal{E}, \bar{a}(\omega), W(\omega))$, es una **variable aleatoria** si cada $a_i(\omega)$ y cada $w_{ij}(\omega)$ son variables aleatorias reales con $\sum_{i=1}^n a_i(\omega) = 1$, $a_i(\omega) > 0$ y $w_{ij}(\omega) \geq 0$, $w_{ii} \equiv 0$ y $w_{ji}(\omega) = w_{ij}(\omega)$.

Si cada una de las $n + \frac{n(n-1)}{2} = \frac{n(n+1)}{2}$ variables aleatorias que definen el grafo $\mathcal{G}(\omega)$ tiene momentos finitos de algún orden, podemos definir los correspondientes momentos para la nueva variable aleatoria \mathcal{G} . Supongamos que cada $w_{ij}(\omega)$ y cada $a_i(\omega)$ tienen momentos finitos de primer orden y si denotamos con $\mathbb{E}(w_{ij})$ y $\mathbb{E}(a_i)$ las correspondientes medias. Entonces, puesto que, integrando para $\omega \in \Omega$ con respecto a $d\mathcal{P}$ la igualdad

$$1 = \sum_{i=1}^n a_i(\omega)$$

obtenemos

$$1 = \sum_{i=1}^n \mathbb{E}(a_i),$$

resulta que tenemos un nuevo grafo de afinidad con atributos que podemos considerar el grafo medio o grafo esperado $\mathbb{E}(\mathcal{G}) = (\mathcal{V}, \mathcal{E}, \mathbb{E}(\bar{a}), \mathbb{E}(W))$, donde $\mathbb{E}(\bar{a})_i = \mathbb{E}(a_i)$ y $(\mathbb{E}(W))_{ij} = \mathbb{E}(w_{ij})$. Si cada una de las $w_{ij}(\omega)$ y $a_i(\omega)$ tiene segundo momento finito, o están en $L^2(\Omega, d\mathcal{P})$, entonces la varianza de cada una de ellas está bien definida

$\sigma_{ij}^2 = \sigma(w_{ij}) = \mathbb{E}(w_{ij} - \mathbb{E}(w_{ij}))^2$ y $\sigma_i^2 = \sigma^2(a_i) = \mathbb{E}(a_i - \mathbb{E}(a_i))^2$. O, de otra manera $\sigma_{ij}^2 = \mathbb{E}(w_{ij}^2) - (\mathbb{E}(w_{ij}))^2$ y $\sigma_i^2 = \mathbb{E}(a_i^2) - (\mathbb{E}(a_i))^2$. Las desviaciones cuadráticas alrededor del grafo esperado $\mathbb{E}(\mathcal{G})$ están regidas por los $\frac{n(n+1)}{2}$ parámetros σ_{ij}^2 y σ_i^2 . Para resumirlo, diremos que $\Sigma^2 = (\sigma_{ij}^2, \sigma_i^2)$ es el vector de varianzas del grafo esperado $\mathbb{E}(\mathcal{G}) = (\mathcal{V}, \mathcal{E}, \mathbb{E}(\bar{a}), \mathbb{E}(W))$.

PROPOSICIÓN 11.4.1. *La función $\mathcal{G} : \Omega \rightarrow \mathbb{G}(\mathcal{V})$ es una variable aleatoria valuada en los grafos de afinidades con atributos sobre \mathcal{V} con la definición precedente si y sólo si \mathcal{G} es una función medible cuando en $\mathbb{G}(\mathcal{V})$ consideramos los Borelianos que define cualquiera de las dos métricas d_F o d_F^{KR} que hemos introducido en el Capítulo 10.*

DEMOSTRACIÓN. Se deduce directamente del hecho que la topología generada por d_F o d_F^{KR} en $\mathbb{G}(\mathcal{V})$ es la misma que la topología en $\mathbb{G}(\mathcal{V})$ que hace de la identificación $\mathbb{G}(\mathcal{V}) \leftrightarrow (\mathbb{R}_0^+)^{\frac{n(n-1)}{2}} \times S$, donde $S = \{\bar{a} = (a_1, \dots, a_n) \in \mathbb{R}^n : a_i \geq 0 \text{ y } \sum_{i=1}^n a_i = 1\} \subset \mathbb{R}^n$, dada por $(\mathcal{V}, \mathcal{E}, \bar{a}, W) \leftrightarrow (w_{12}, w_{13}, \dots, w_{1n}; w_{23}, w_{24}, \dots, w_{2n}; \dots; w_{n-1n}; a_1, \dots, a_n)$, un homeomorfismo. \square

11.5. Aleatorización de grafos clásicos

Con el procedimiento indicado para el grafo bipartito de la Sección 10.4, tomando un número aleatorio $a_i(\omega)$ uniformemente distribuido en el intervalo $(\frac{9}{10}, \frac{11}{10})$ con media $a_i = 1$ para $i = 1, \dots, 9$, y números aleatorios en el intervalo $(\frac{9}{10}, \frac{11}{10})$ alrededor de w_{ij} cuando $w_{ij} = 1$ y en el intervalo $(0, \frac{1}{10})$ cuando $w_{ij} = 0$, en la tabla de los Cuadros 1 y 2 a modo ilustrativo vamos a comparar los autovalores del Laplaciano de los grafos clásicos (considerados en la Sección 1.4) con los autovalores del Laplaciano aleatorio generado por los atributos $\bar{a}(\omega)$ y las afinidades $W(\omega)$ de cada grafo. En las gráficas se dibujan 20 muestras. En cada nodo de la columna de “Esquema” de los Cuadros 1 y 2, etiquetamos sólo uno de estos ensayos aleatorios. Para describir gráficamente la perturbación de la matriz W de unos y ceros de cada grafo, usamos líneas rojas para dibujar las aristas cuyo peso \tilde{w}_{ij} es próximo a uno y usamos líneas de puntos celestes para representar las perturbaciones cercanas a cero. En el sentido estricto todos los esquemas perturbados son grafos completos pero las ponderaciones de sus aristas, todavía dejan ver el grafo original que ha sido perturbado aleatoriamente.

| Grafo | Esquema | Autovalores |
|----------|---------|-------------|
| Camino | | |
| Estrella | | |
| Ciclo | | |

CUADRO 1. Aleatoriedad de autovalores alrededor de grafos clásicos.

| Grafo | Esquema | Autovalores |
|-----------|---------|-------------|
| Rueda | | |
| Bipartito | | |
| Completo | | |

CUADRO 2. Aleatoriedad de autovalores alrededor de grafos clásicos.

11.6. Aleatorización de afinidad y atributos en el grafo de SUBE

Aplicamos ahora un esquema similar al caso de los grafos de AMBA con aleatorización de afinidades y de atributos de infectados por distritos. Procedemos a describir brevemente el procedimiento y el algoritmo que Python ejecuta para proveernos dos gráficas contenidas en las Figuras 11.6.3 y 11.6.4. La aleatorización de las matrices de SUBE está realizada uniformemente y proporcionalmente a la intensidad del transporte entre cada par de distritos para los meses de abril y junio, precisamente, para $i < j$, $(w_A)_{ij}(\omega)$ es una variable aleatoria uniformemente distribuida en el intervalo $(\frac{9}{10}(w_A)_{ij}, \frac{11}{10}(w_A)_{ij})$. Para $i > j$ tomamos, $(w_A)_{ij}(\omega) = (w_A)_{ji}(\omega)$ y $(w_J)_{ij}(\omega) = (w_J)_{ji}(\omega)$. Para los vectores de infectados procedemos de manera similar tomando $(a_A^I)_i(\omega)$ uniformemente distribuido en el intervalo $(\frac{9}{10}(a_A^I)_i, \frac{11}{10}(a_A^I)_i)$. Análogamente para $(w_J)_{ij}(\omega)$ y para el vector \bar{a}_J^I . Para cada uno de 10 ensayos, luego de normalizar las matrices aleatorias y los vectores aleatorios a probabilidad (suma uno), calculamos el espectro de cada uno de los 10 operadores de Laplace que quedan definidos y graficamos los correspondientes autovalores.

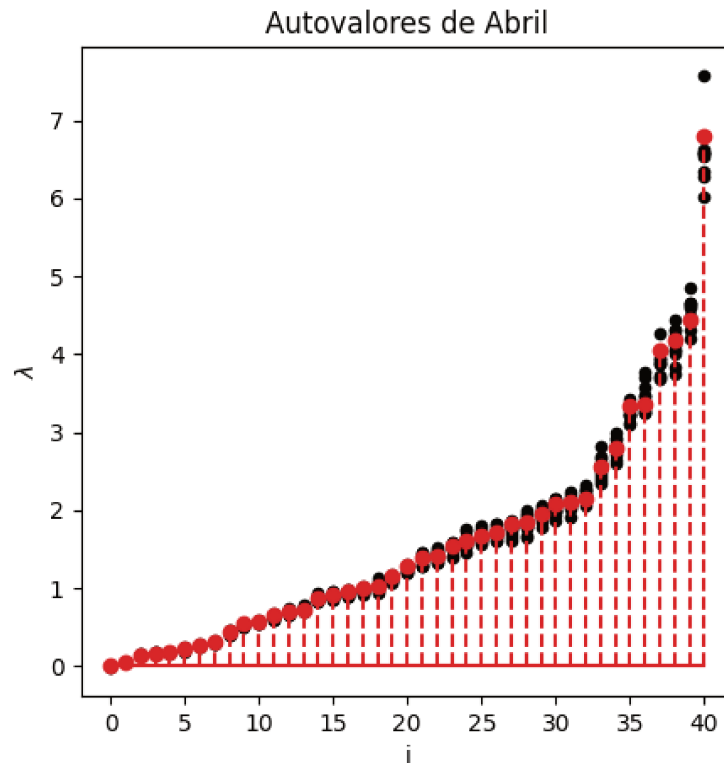


FIGURA 11.6.3. Aleatoriedad de autovalores para SUBE en AMBA con atributos de infectados COVID-19 en Abril de 2020.

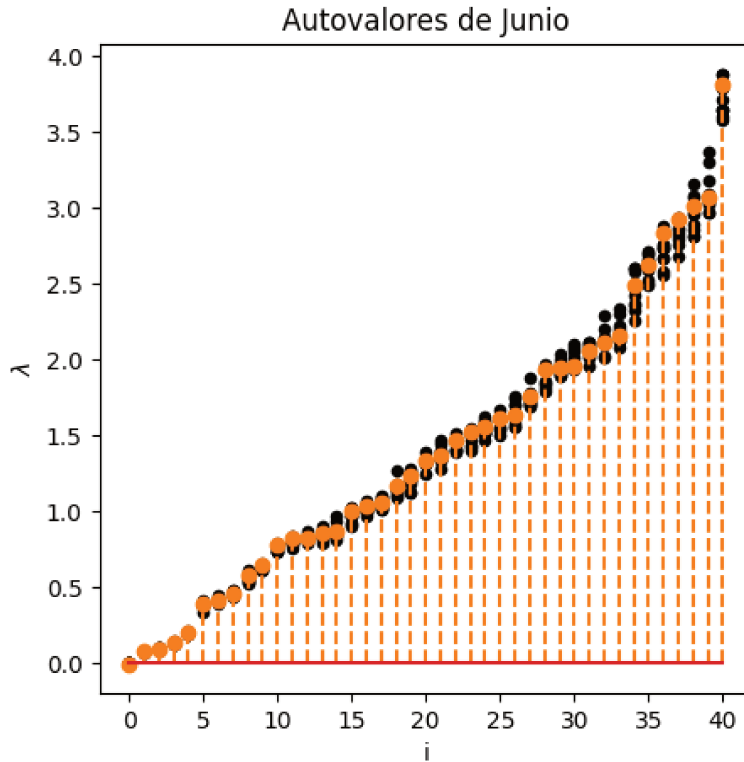
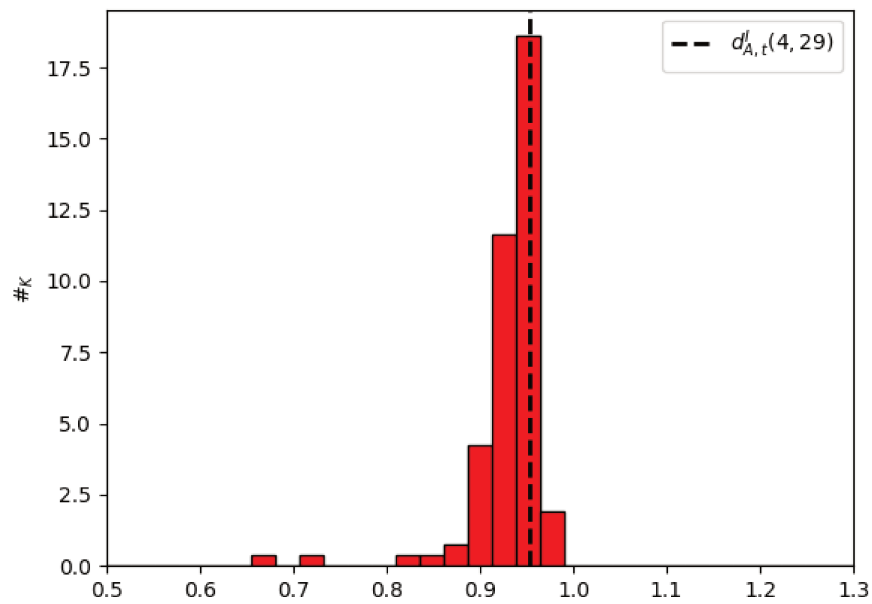
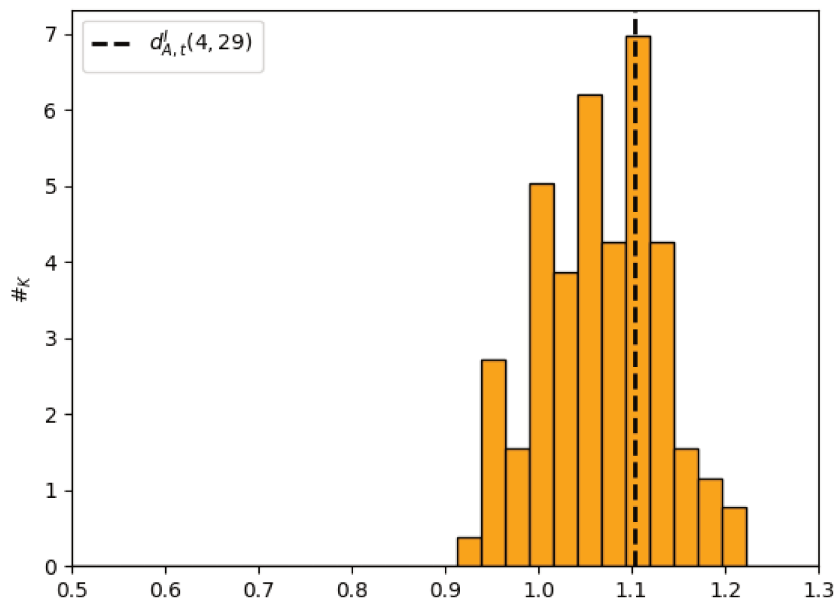


FIGURA 11.6.4. Aleatoriedad de autovalores para SUBE en AMBA con atributos de infectados COVID-19 en Junio de 2020.

En el capítulo próximo demostraremos que el empirismo precedente se sustenta en desigualdades de concentración que estudiaremos para distribuciones subgaussianas que incluyen a las distribuciones uniformes que hemos usado en estos experimentos. Del mismo modo, y usando los cálculos anteriores podemos también obtener las distancias difusivas aleatorias correspondientes a dos distritos cualesquiera, por ejemplo, como antes Ezeiza (4) y CABA (29). Para 100 ensayos tenemos las situaciones correspondientes a Abril y Junio para $t = 0.01$ que se muestran en las Figuras 11.6.5 y 11.6.6. La línea vertical (punteada) corresponde a los valores medios de $d_{A,t}^I(4, 29)$ y de $d_{J,t}^I(4, 29)$ respectivamente. El histograma registra el número $\#_k$ de ensayos, de los 100, para los cuales $d_{A,t}^I(\omega)(4, 29)$ y $d_{J,t}^I(\omega)(4, 29)$ caen en el intervalo $I_k = [\frac{k}{40}, \frac{k+1}{40})$ con $k = 20, 21, \dots, 51$.

FIGURA 11.6.5. Histograma de $d_{A,t}^I(\omega)(4, 29)$ para 100 ensayos.FIGURA 11.6.6. Histograma de $d_{J,t}^I(\omega)(4, 29)$ para 100 ensayos.

Concentración de variables aleatorias valuadas en grafos métricos y de afinidades con atributos. Concentración espectral

12.1. Introducción

En este capítulo nos proponemos hacer un análisis de las grandes desviaciones alrededor de las medias para los objetos aleatorios que quedan definidos por variables aleatorias con valores en $\mathcal{M}^l(X_0, d_0)$ como lo son $d_{GL}^{KR}((X(\omega), d(\omega), \mu(\omega)), (\bar{X}, \bar{d}, \bar{\mu}))$ y en el caso de grafos métricos y de afinidades, los autovalores $\lambda_j(\omega)$ de los Laplacianos aleatorios generados. Usaremos el caso subgaussiano de la teoría de concentración de Cramér-Chernoff. El capítulo está organizado en cinco secciones. En las Secciones 12.2 y 12.3 exponemos los aspectos básicos del método de Cramér-Chernoff en el caso subgaussiano. En la Sección 12.4 consideramos la concentración de $d_{GL}^{KR}(\mathcal{G}(\omega), \mathcal{G}_0)$ siendo \mathcal{G}_0 el grafo medio de los $\mathcal{G}(\omega)$. En la Sección 12.5 estudiamos concentración del Laplaciano y del espectro y también un caso particular de concentración de las métricas difusivas asociadas.

12.2. El método de Cramér-Chernoff. Variables aleatorias subgaussianas. Concentración

Aunque adaptaremos notación y contexto a nuestro desarrollo general, esta sección sigue los lineamientos expuesto en [BLM13]. El punto de partida del análisis del decaimiento de la distribución de una variable aleatoria en el infinito, cuando la variable tiene momentos finitos, es la desigualdad de Markov y su más conocida consecuencia la desigualdad de Chebychev.

Todas las desigualdades que interesan en este capítulo buscan cuantificar o estimar superiormente grandes desviaciones alrededor de la media. Esto consiste esencialmente en la estimación del tamaño de las “colas” de las distribuciones de probabilidad. Cuando la variable aleatoria X es clásica, que toma valores reales, hay dos colas que estimar: $X - \mathbb{E}(X) \geq t, t > 0$ y $X - \mathbb{E}(X) \leq -t, t > 0$. Aquí usamos \mathbb{E} para denotar la esperanza. Por otra parte para variables aleatorias valuadas en espacios más generales como \mathbb{R}^n o

mucho más generales del tipo \mathcal{M} considerados en el Capítulo 11, nos interesará estimar probabilidades de grandes desviaciones de la distancia d_{GL}^{KR} a algún punto particular del espacio, que puede ser entendido como una media. Más precisamente, si X es una variable aleatoria valuada en un espacio métrico con métrica d , nos interesará estimar superiormente

$$\mathcal{P}(\{\omega \in \Omega : d(X(\omega), X_0) \geq t\})$$

para t grande y X_0 fijo en el espacio métrico, bajo ciertas condiciones de momentos en $X(\omega)$. El resultado básico en el que se fundamenta el método de Cramér-Chernoff es el siguiente lema de Markov.

LEMA 12.2.1. *Sea $(\Omega, \mathcal{F}, \mathcal{P})$ un espacio de probabilidad e $Y : \Omega \rightarrow \mathbb{R}$ una variable aleatoria, entonces para todo $t > 0$ valen las desigualdades*

- a) $\mathcal{P}(\{\omega \in \Omega : Y(\omega) \geq t\}) \leq \frac{1}{t} \int_{\{Y(\omega) \geq t\}} Y(\omega) d\mathcal{P}(\omega);$
 b) $\mathcal{P}(\{\omega \in \Omega : Y(\omega) \leq -t\}) \leq \frac{1}{t} \int_{\{Y(\omega) \leq -t\}} (-Y(\omega)) d\mathcal{P}(\omega).$

DEMOSTRACIÓN. a) $\mathcal{P}(\{Y \geq t\}) = \int_{\{Y \geq t\}} d\mathcal{P} = \int_{\{\frac{Y}{t} \geq 1\}} d\mathcal{P} \leq \frac{1}{t} \int_{\{Y \geq t\}} Y d\mathcal{P}.$

b) sigue de a) tomando $-Y$ en vez de Y . □

Observamos que, aunque Y no pertenezca a $L^1(\Omega, \mathcal{F}, d\mathcal{P})$, las desigualdades precedentes son válidas.

Cuando una variable aleatoria real toma sólo valores no negativos (respectivamente no positivos) la desigualdad en a) provee información, la desigualdad en b) es trivial ya que el lado izquierdo es nulo. En particular cuando la variable aleatoria tiene suficientes momentos finitos, tenemos desigualdades de Chebychev y sus versiones para $p \neq 2$.

COROLARIO 12.2.2. *Sea X una variable aleatoria en $L^2(\Omega, \mathcal{F}, \mathcal{P})$. Entonces*

$$\mathcal{P}(\{|X - \mathbb{E}(X)| \geq t\}) \leq \frac{1}{t^2} \int_{\Omega} |X - \mathbb{E}(X)|^2 d\mathcal{P} = \frac{\sigma^2}{t^2}.$$

DEMOSTRACIÓN. Tomando $Y(\omega) = |X(\omega) - \mathbb{E}(X)|^2$ en a) del Lema 12.2.1, tenemos que

$$\begin{aligned} \mathcal{P}(\{|X - \mathbb{E}(X)| \geq t\}) &= \mathcal{P}(\{|X - \mathbb{E}(X)|^2 \geq t^2\}) \\ &\leq \frac{1}{t^2} \int_{\{|X - \mathbb{E}(X)|^2 \geq t^2\}} |X - \mathbb{E}(X)|^2 d\mathcal{P} \end{aligned}$$

$$\begin{aligned} &\leq \frac{1}{t^2} \int_{\Omega} |X - \mathbb{E}(X)|^2 d\mathcal{P} \\ &= \frac{\sigma^2}{t^2}. \end{aligned}$$

□

PROPOSICIÓN 12.2.3. *Sea X una variable aleatoria tal que $X(\Omega) \subseteq I$, con I un intervalo, una semirrecta o toda la recta real. Sea $\phi : I \rightarrow \mathbb{R}_0^+$ no decreciente. Entonces la desigualdad*

$$\mathcal{P}(\{\omega \in \Omega : X(\omega) \geq t\}) \leq \frac{\mathbb{E}(\phi(X))}{\phi(t)}$$

vale para todo $t \in I$ tal que $\phi(t) > 0$.

DEMOSTRACIÓN. Por la monotonía de ϕ , usando la desigualdad de Markov en el Lema 12.2.1 parte a)

$$\begin{aligned} \mathcal{P}(\{X \geq t\}) &\leq \mathcal{P}(\{\phi(X) \geq \phi(t)\}) \\ &\leq \frac{1}{\phi(t)} \int_{\{\phi(X) \geq \phi(t)\}} \phi(X) d\mathcal{P} \\ &\leq \frac{1}{\phi(t)} \int_{\Omega} \phi(X) d\mathcal{P} \\ &= \frac{\mathbb{E}(\phi(X))}{\phi(t)}. \end{aligned}$$

□

El siguiente resultado contiene el caso particular de la Proposición 12.2.3 que usaremos en el método de Cramér-Chernoff en el que tomamos una familia de funciones ϕ .

PROPOSICIÓN 12.2.4. *Sea X una variable aleatoria con esperanza finita. Sea $\lambda > 0$ e $I = \mathbb{R}$. Entonces*

$$(12.1) \quad \mathcal{P}(\{\omega \in \Omega : X(\omega) - \mathbb{E}(X) \geq t\}) \leq \frac{\mathbb{E}(e^{\lambda(X - \mathbb{E}(X))})}{e^{\lambda t}},$$

para todo $t \in \mathbb{R}$.

DEMOSTRACIÓN. Notamos que $X - \mathbb{E}(X)$ es una nueva variable aleatoria que toma también valores en \mathbb{R} , aplicamos a ella la Proposición 12.2.3 con $\phi_{\lambda}(t) = e^{\lambda t}$. □

Observamos que, con la sola hipótesis de $X \in L^1(\Omega, \mathcal{F}, \mathcal{P})$, no está garantizada la finitud de $\mathbb{E}(e^{\lambda(X - \mathbb{E}(X))})$, pero la desigualdad es trivial en el caso en que esto no ocurra.

Cuando tenemos variables aleatorias generales valuadas en un espacio métrico, también podemos probar desigualdades de tipos Markov y Chebychev como las anteriores para las distancias de los ensayos aleatorios a algún punto fijo del espacio. La siguiente proposición contiene un resultado en ese sentido.

PROPOSICIÓN 12.2.5. *Sea $I = \mathbb{R}_0^+$, $\phi : I \rightarrow \mathbb{R}_0^+$ no decreciente, (Θ, η) un espacio métrico (o casi-métrico) y $X : \Omega \rightarrow (\Theta, \eta)$ una variable aleatoria valuada en (Θ, η) . Entonces, si θ_0 es un punto cualquiera en Θ , tenemos*

$$\mathcal{P}(\{\omega \in \Omega : \eta(X(\omega), \theta_0) \geq t\}) \leq \frac{\mathbb{E}(\phi(\eta(X, \theta_0)))}{\phi(t)},$$

para todo $t > 0$.

DEMOSTRACIÓN. Aplicamos la desigualdad de Markov a la variable aleatoria $\phi(\eta(X(\omega), \theta_0))$ con $\phi(t)$ en lugar de t . □

Observamos que las estimaciones (12.1) provistas por la Proposición 12.2.4 son válidas para todo $\lambda > 0$, aunque el miembro izquierdo de (12.1) no depende de λ . El método Cramér-Chernoff consiste precisamente en encontrar el parámetro λ^* que minimiza el miembro derecho de (12.1) con el objeto de obtener la estimación superior más justa para $\mathcal{P}(\{X - \mathbb{E}(X) \geq t\})$. Por simplicidad notacional, escribiremos $Y(\omega) = X(\omega) - \mathbb{E}(X)$.

La **función generadora de momentos** de la variable aleatoria Y es la función $\Gamma_Y : \mathbb{R}^+ \rightarrow \mathbb{R}^+ \cup \{+\infty\}$ definida por $\Gamma_Y(\lambda) = \mathbb{E}(e^{\lambda Y}) = \int_{\Omega} e^{\lambda Y(\omega)} d\mathcal{P}(\omega)$. Denotamos con Ψ_Y al logaritmo de la función generadora de momentos, $\Psi_Y(\lambda) = \log \Gamma_Y(\lambda) = \log \mathbb{E}(e^{\lambda Y})$. Notar que $\Psi_Y : \mathbb{R}^+ \rightarrow \mathbb{R} \cup \{+\infty\}$. Consideraremos, para ilustrar el método el caso normal en el que la variable aleatoria Y tiene media nula, varianza σ^2 y está normalmente distribuida. En otros términos $Y \sim N(0, \sigma^2)$.

PROPOSICIÓN 12.2.6. *Sea $Y \sim N(0, \sigma^2)$. Entonces*

- a) $\Gamma_Y(\lambda) = e^{\frac{\lambda^2 \sigma^2}{2}}$;
- b) $\Psi_Y(\lambda) = \frac{\lambda^2 \sigma^2}{2}$;
- c) $\log \mathcal{P}(\{Y \geq t\}) \leq \log(e^{-\lambda t} \Gamma_Y(\lambda)) = -\lambda t + \Psi_Y(\lambda) = -(\lambda t - \frac{\lambda^2 \sigma^2}{2})$;
- d) $\Psi_Y^*(t) := \sup_{\lambda > 0} (\lambda t - \Psi_Y(\lambda)) = \sup_{\lambda > 0} (\lambda t - \frac{\lambda^2 \sigma^2}{2}) = \frac{t^2}{2\sigma^2}$;

e) $\mathcal{P}(\{Y \geq t\}) \leq e^{-\frac{t^2}{2\sigma^2}}$ para $t > 0$.

DEMOSTRACIÓN. a)

$$\begin{aligned}\Gamma_Y(\lambda) &= \mathbb{E}(e^{\lambda Y}) = \frac{1}{\sqrt{2\pi}\sigma} \int_{y \in \mathbb{R}} e^{\lambda y} e^{-\frac{y^2}{2\sigma^2}} dy \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{y \in \mathbb{R}} e^{-\frac{1}{2\sigma^2}(y^2 - 2\lambda\sigma^2 + \lambda^2\sigma^4)} e^{\frac{1}{2}\lambda^2\sigma^2} dy \\ &= e^{\frac{\lambda^2\sigma^2}{2}} \frac{1}{\sqrt{2\pi}\sigma} \int_{y \in \mathbb{R}} e^{-\frac{1}{2\sigma^2}(y - \lambda\sigma^2)^2} dy \\ &= e^{\frac{\lambda^2\sigma^2}{2}}.\end{aligned}$$

b) sigue de a) tomando logaritmos. c) sigue de la Proposición 12.2.4 tomando logaritmos. Veamos d). Sea $t \in \mathbb{R}$ fijo, si $\eta_t(\lambda) = \lambda t - \frac{\lambda^2\sigma^2}{2}$, $\frac{d\eta_t}{d\lambda} = t - 2\frac{\lambda\sigma^2}{2} = t - \lambda\sigma^2$, que se anula en $\lambda = \frac{t}{\sigma^2}$, donde η_t alcanza su máximo, cuyo valor es

$$\begin{aligned}\eta_t\left(\frac{t}{\sigma^2}\right) &= \frac{t}{\sigma^2}t - \left(\frac{t}{\sigma^2}\right)^2 \frac{\sigma^2}{2} \\ &= \frac{t^2}{\sigma^2} - \frac{t^2}{2\sigma^2} \\ &= \frac{t^2}{2\sigma^2}.\end{aligned}$$

e) por d) $\log \mathcal{P}(\{Y \geq t\}) \leq -\frac{t^2}{2\sigma^2}$ entonces $\mathcal{P}(\{Y \geq t\}) \leq e^{-\frac{t^2}{2\sigma^2}}$. \square

12.3. Variables aleatorias subgaussianas

La función $\Psi_Y^*(t)$ está definida en casos muchos más generales que el gaussiano como

$$\Psi_Y^*(t) = \sup_{\lambda > 0} (\lambda t - \Psi_Y(\lambda))$$

que puede verse como un caso particular de la transformación de Legendre de Ψ_Y cuando ésta es convexa.

El carácter gaussiano en las variables aleatorias consideradas en la Proposición 12.2.6 queda determinado por la identidad del inciso b) de la proposición: $\Psi_Y(\lambda) = \frac{\lambda^2\sigma^2}{2}$ y justamente esta identidad es la que determina la estimación para las grandes desviaciones de la media si seguimos tomando $Y = X - \mathbb{E}(X)$ en e) de la misma proposición: $\mathcal{P}(\{Y \geq t\}) \leq e^{-\frac{t^2}{2\sigma^2}}$. Muchas otras clases de variables aleatorias no gaussianas, como las que tienen distribuciones uniformes en intervalos acotados de \mathbb{R} , admiten una dominación gaussiana.

Atendiendo a los ejemplos y aplicaciones de los capítulos precedentes en los que hemos usado esencialmente distribuciones uniformes para modelizar los datos, limitaremos el análisis a **variables aleatorias subgaussianas**. Diremos que una variable aleatoria Y con media nula es **subgaussiana** con factor de varianza $\nu > 0$ si

$$\Psi_Y(\lambda) = \log \Gamma_Y(\lambda) = \log \mathbb{E}(e^{\lambda Y}) \leq \frac{\lambda^2 \nu}{2}$$

para todo $\lambda \in \mathbb{R}$. Claro, por *b*) de la Proposición 12.2.6, que toda variable aleatoria gaussiana es subgaussiana con $\nu = \sigma^2$. Menos claro, pero igualmente cierto, es que las variables aleatorias que toman valores en conjuntos acotados de \mathbb{R} son subgaussianas. Este resultado es debido a Hoeffding [Hoe63], y por completitud lo probamos a continuación. En primer lugar probamos un lema simple de reescalado de variables aleatorias con valores acotados.

LEMA 12.3.1.

- a)* Si Z es una variable aleatoria que toma valores en el intervalo $[0, r]$, $r > 0$, entonces $\sigma^2(Z) \leq \frac{r^2}{4}$.
- b)* Si Y es una variable aleatoria que toma valores en el intervalo $[a, b]$ con $-\infty < a < b < +\infty$, entonces $\sigma^2(Y) \leq \frac{(b-a)^2}{4}$.

DEMOSTRACIÓN. *a)* Notamos que la función cuadrática $\phi(x) = x(r - x)$ tiene un máximo en $x_M = \frac{r}{2}$. Calculemos y estimemos la varianza de Z ,

$$\begin{aligned} \sigma^2(Z) &= \mathbb{E}(Z^2) - (\mathbb{E}(Z))^2 \\ &= \int_{\Omega} Z^2 d\mathcal{P} - (\mathbb{E}(Z))^2 \leq r\mathbb{E}(Z) - (\mathbb{E}(Z))^2 \\ &= \mathbb{E}(Z)(r - \mathbb{E}(Z)) = \phi(\mathbb{E}(Z)) \\ &\leq \phi(X_M) = \frac{r^2}{4}. \end{aligned}$$

b) Dado que $a \leq Y(\omega) \leq b$ para casi todo ω , también tenemos que $0 \leq Y(\omega) - a \leq b - a$. Sea $Z(\omega) = Y(\omega) - a$, aplicamos *a)* a $Z - a$ con $r = b - a$ y obtenemos

$$\sigma^2(Z) = \sigma^2(Z - a) \leq \frac{(b - a)^2}{4}.$$

□

TEOREMA 12.3.2. *Sea $X : \Omega \rightarrow [a, b]$ una variable aleatoria con $-\infty < a < b < +\infty$ fijos. Entonces $X - \mathbb{E}(X)$ es subgaussiana con factor de varianza $\frac{(b-a)^2}{4}$.*

DEMOSTRACIÓN. Tomando $Y = X - \mathbb{E}(X)$, $\alpha = a - \mathbb{E}(X)$ y $\beta = b - \mathbb{E}(X)$ basta demostrar que Y es subgaussiana con factor de varianza $\frac{(\beta-\alpha)^2}{4} = \frac{(b-a)^2}{4}$. Recordemos que $\Psi_Y(\lambda) = \log \mathbb{E}(e^{\lambda Y})$. Sea μ la medida de probabilidad que distribuye a Y . Es decir $\mu(A) = \mathcal{P}(Y^{-1}(A))$ para A Boreliano en \mathbb{R} . Para cada $\lambda \in \mathbb{R}$ definimos una nueva distribución de probabilidad μ_λ dada por $\mu_\lambda(A) = \int_A e^{-\Psi_Y(\lambda)} e^{\lambda x} d\mu(x) = e^{-\Psi_Y(\lambda)} \int_A e^{\lambda x} d\mu(x) = \frac{1}{\mathbb{E}(e^{\lambda Y})} \int_A e^{\lambda x} d\mu(x)$. Notemos que si $A \cap [\alpha, \beta] = \emptyset$, entonces $\mu_\lambda(A) = \mu(A) = 0$. En otras palabras, los soportes de todas las μ_λ están contenidas en $[\alpha, \beta]$. Notemos también que en particular, cuando $A = \mathbb{R}$, tenemos que $\mu_\lambda(\mathbb{R}) = \frac{1}{\mathbb{E}(e^{\lambda Y})} \int_{\mathbb{R}} e^{\lambda x} d\mu(x) = \frac{\mathbb{E}(e^{\lambda Y})}{\mathbb{E}(e^{\lambda Y})} = 1$. Por el Lema 12.3.1, tenemos también que, cualquier Z cuya distribución sea μ_λ tiene varianza acotada por $\frac{(\beta-\alpha)^2}{4}$, $\sigma^2(Z) \leq \frac{(\beta-\alpha)^2}{4}$. Observamos que la función $\Psi_Y(\lambda)$ es suave como función de λ . Vamos a estudiar el comportamiento de $\Psi_Y(\lambda)$ usando el teorema de Taylor para probar la tesis. Notemos primero que $\Psi_Y(0) = \log \mathbb{E}(e^{0Y}) = \log 1 = 0$. Calculemos ahora $\frac{d\Psi_Y}{d\lambda}$;

$$\begin{aligned} \frac{d\Psi_Y}{d\lambda} &= \frac{d}{d\lambda} \log \mathbb{E}(e^{\lambda Y}) \\ &= \frac{1}{\mathbb{E}(e^{\lambda Y})} \int_{\Omega} Y e^{\lambda Y} d\mathcal{P} \\ &= \frac{\mathbb{E}(Y e^{\lambda Y})}{\mathbb{E}(e^{\lambda Y})} \\ &= e^{-\Psi_Y(\lambda)} \mathbb{E}(Y e^{\lambda Y}). \end{aligned}$$

Entonces, $\frac{d\Psi_Y}{d\lambda}(0) = \mathbb{E}(Y e^{0Y}) = \mathbb{E}(Y) = 0$. Calculemos ahora la segunda derivada de $\Psi_Y(\lambda)$;

$$\begin{aligned} \frac{d^2\Psi_Y}{d\lambda^2} &= \frac{d}{d\lambda} [e^{-\Psi_Y(\lambda)} \mathbb{E}(Y e^{\lambda Y})] \\ &= -e^{-\Psi_Y(\lambda)} \frac{d\Psi_Y}{d\lambda} \mathbb{E}(Y e^{\lambda Y}) + e^{-\Psi_Y(\lambda)} \mathbb{E}(Y^2 e^{\lambda Y}) \\ &= -e^{-2\Psi_Y(\lambda)} [\mathbb{E}(Y e^{\lambda Y})]^2 + e^{-\Psi_Y(\lambda)} \mathbb{E}(Y^2 e^{\lambda Y}) \\ &= \mathbb{E}(Y^2 e^{-\Psi_Y(\lambda)} e^{\lambda Y}) - [\mathbb{E}(Y e^{-\Psi_Y(\lambda)} e^{\lambda Y})]^2. \end{aligned}$$

Sea Z una variable aleatoria tal que su distribución es μ_λ , es decir

$$\mathcal{P}(\{Z \in A\}) = \int_A d\mu_\lambda(x) = \mu_\lambda(A) = \int_A e^{-\Psi_Y(\lambda)} e^{\lambda x} d\mu(x).$$

Entonces, como μ es la distribución de Y , tenemos

$$\mathbb{E}(Z) = \int_{\mathbb{R}} z e^{-\Psi_Y(\lambda)} e^{\lambda z} d\mu(z) = \mathbb{E}(Y e^{-\Psi_Y(\lambda)} e^{\lambda Y})$$

y

$$\sigma^2(Z) = \mathbb{E}(Z^2) - [\mathbb{E}(Z)]^2 = \mathbb{E}(Y^2 e^{-\Psi_Y(\lambda)} e^{\lambda Y}) - [\mathbb{E}(Y e^{-\Psi_Y(\lambda)} e^{\lambda Y})]^2 = \frac{d^2 \Psi_Y}{d\lambda^2}.$$

Entonces

$$\frac{d^2 \Psi_Y}{d\lambda^2}(\lambda) = \sigma^2(Z) \leq \frac{(\beta - \alpha)^2}{4}.$$

Con estos tres datos, $\Psi_Y(0) = 0 = \frac{d\Psi_Y}{d\lambda}(0)$ y $\frac{d^2 \Psi_Y}{d\lambda^2} \leq \frac{(\beta - \alpha)^2}{4}$, vamos a Taylor de segundo orden para Ψ_Y alrededor del origen $\lambda_0 = 0$,

$$\begin{aligned} \Psi_Y(\lambda) &= \Psi_Y(0) + \lambda \frac{d\Psi_Y}{d\lambda}(0) + \frac{\lambda^2}{2} \frac{d^2 \Psi_Y}{d\lambda^2}(\xi) \\ &= \frac{\lambda^2}{2} \frac{d^2 \Psi_Y}{d\lambda^2}(\xi) \\ &\leq \frac{\lambda^2 (\beta - \alpha)^2}{8} \\ &= \frac{\lambda^2}{2} \frac{(\beta - \alpha)^2}{4}, \end{aligned}$$

entonces Y es subgaussiana con factor de varianza $\frac{(\beta - \alpha)^2}{4}$. □

El siguiente resultado extiende la Proposición 12.2.6 a variables subgaussianas.

PROPOSICIÓN 12.3.3. *Sea Y una variable aleatoria con media nula que es subgaussiana con factor de varianza ν . Entonces la desigualdad*

$$\mathcal{P}(\{Y \geq t\}) \leq e^{-\frac{t^2}{2\nu}}$$

vale para todo $t > 0$.

DEMOSTRACIÓN. Por hipótesis $\log \mathbb{E}(e^{\lambda Y}) = \Psi_Y(\lambda) \leq \frac{\lambda^2 \nu}{2}$. Entonces, por la Proposición 12.2.4

$$\log \mathcal{P}(\{Y \geq t\}) \leq \log [e^{-\lambda t} \mathbb{E}(e^{\lambda Y})]$$

$$\begin{aligned}
&= -\lambda t + \log \mathbb{E}(e^{\lambda Y}) \\
&= -\lambda t + \Psi_Y(\lambda) \\
&\leq -\lambda t + \frac{\lambda^2 \nu}{2}
\end{aligned}$$

para todo λ . Entonces

$$\log \mathcal{P}(\{Y \geq t\}) \leq -\lambda^* + \frac{(\lambda^*)^2 \nu}{2},$$

donde λ^* es el parámetro que minimiza la función $-(\lambda t - \frac{\lambda^2 \nu}{2})$ para t fijo. Es decir, λ^* es el que maximiza la función cuadrática de λ dada por $\lambda t - \frac{\lambda^2 \nu}{2}$. Ese máximo se realiza en $\lambda^* = \frac{t}{\nu}$ y toma el valor $\frac{t}{\nu} t - \frac{t^2 \nu}{\nu^2 2} = \frac{1}{2} \frac{t^2}{\nu}$. Por consiguiente $\log \mathcal{P}(\{Y \geq t\}) \leq -\frac{1}{2} \frac{t^2}{\nu}$. De aquí que

$$\mathcal{P}(\{Y \geq t\}) \leq e^{-\frac{t^2}{2\nu}}.$$

□

COROLARIO 12.3.4. *Sea Y una variable aleatoria que toma valores en el intervalo $[a, b]$ con $-\infty < a < b < +\infty$. Entonces*

$$\mathcal{P}(\{Y - \mathbb{E}(Y) \geq t\}) \leq e^{-\frac{2t^2}{(b-a)^2}}.$$

DEMOSTRACIÓN. Sigue la proposición anterior y del teorema de Hoeffding. □

De manera análoga se prueba que

$$\mathcal{P}(\{Y - \mathbb{E}(Y) \leq -t\}) \leq e^{-\frac{2t^2}{(b-a)^2}}.$$

12.4. Concentración de grafos con métricas y atributos aleatorios alrededor del grafo medio

El contexto será el del Teorema 8.3.2. Estaremos tomando puntos $\bar{\xi}$ y vectores \bar{a} aleatorios alrededor de valores prescriptos $\bar{\xi}^0$ y $\bar{a}^0 = (a_1^0, \dots, a_n^0)$, y nos proponemos estudiar la concentración de $(\mathcal{V}, d_{\bar{\xi}(\omega)}, \mu_{\bar{a}(\omega)})$ alrededor de $(\mathcal{V}, d_{\bar{\xi}^0}, \mu_{\bar{a}^0})$ con respecto a la métrica d_{GL}^{KR} .

TEOREMA 12.4.1. *Sea $(\Omega, \mathcal{F}, \mathcal{P})$ un espacio de probabilidad. Sean $\xi_i : \Omega \rightarrow B_\rho(\xi_i^0, \frac{R}{2n})$, $i \in \mathcal{V}$, n -variables aleatorias independientes valuadas en Ξ . Sean $a_i : \Omega \rightarrow (a_i^0 - m, a_i^0 + m)$, con $m = \min_{i \in \mathcal{V}} a_i^0$, n variables aleatorias independientes. Sea $\mathcal{G}_0 = (\mathcal{V}, d_{\bar{\xi}^0}, \mu_{\bar{a}^0})$ y para cada*

$\omega \in \Omega$, sea $\mathcal{G}_\omega = (\mathcal{V}, d_{\bar{\xi}(\omega)}, \mu_{\bar{a}(\omega)})$. Entonces, si Θ es la constante en el Teorema 8.3.2 tenemos que

$$\mathcal{P}(\{\omega \in \Omega : \mathfrak{d}_{GL}^{KR}(\mathcal{G}_\omega, \mathcal{G}_0) > t\}) \leq e^{-\frac{1}{2} \frac{n^3}{\Theta^4 R^4} t^4} + e^{-\frac{1}{32} \frac{1}{nm^4 \Theta^4} t^4},$$

para $t > \Theta \max \left\{ \frac{R}{\sqrt{n}}, 2m\sqrt{n} \right\}$.

DEMOSTRACIÓN. Por el Teorema 8.3.2 y las restricciones en los valores de las variables aleatorias ξ_i y a_i tenemos, para todo $\omega \in \Omega$, que

$$[\mathfrak{d}_{GL}^{KR}(\mathcal{G}_\omega, \mathcal{G}_0)]^2 \leq \Theta^2 \left(\sum_{i=1}^n \rho^2(\xi_i(\omega), \xi_i^0) + \sum_{i=1}^n |a_i(\omega) - a_i^0|^2 \right),$$

con Θ la del Teorema 8.3.2. Entonces,

$$\begin{aligned} \mathcal{P}(\{\mathfrak{d}_{GL}^{KR}(\mathcal{G}_\omega, \mathcal{G}_0) > t\}) &= \mathcal{P}(\{\mathfrak{d}_{GL}^{KR}(\mathcal{G}_\omega, \mathcal{G}_0)^2 > t^2\}) \\ &\leq \mathcal{P} \left(\left\{ \Theta^2 \sum_{i=1}^n \rho^2(\xi_i(\omega), \xi_i^0) \geq \frac{t^2}{2} \right\} \right) + \mathcal{P} \left(\left\{ \Theta^2 \sum_{i=1}^n |a_i(\omega) - a_i^0|^2 \geq \frac{t^2}{2} \right\} \right) \\ &= I(t) + II(t). \end{aligned}$$

Comenzamos estimando el primer sumando del miembro derecho de la desigualdad precedente, usando la independencia y el carácter subgaussiano de las variables aleatorias $\rho^2(\xi_i(\omega), \xi_i^0)$, provisto por el Lema de Hoeffding. En efecto

$$\begin{aligned} I(t) &= \mathcal{P} \left(\left\{ \sum_{i=1}^n [\rho^2(\xi_i(\omega), \xi_i^0) - \mathbb{E}(\rho^2(\xi_i, \xi_i^0))] + \sum_{i=1}^n \mathbb{E}(\rho^2(\xi_i, \xi_i^0)) \geq \frac{t^2}{2\Theta^2} \right\} \right) \\ &\leq \mathcal{P} \left(\left\{ \sum_{i=1}^n [\rho^2(\xi_i(\omega), \xi_i^0) - \mathbb{E}(\rho^2(\xi_i, \xi_i^0))] \geq \frac{t^2}{4\Theta^2} \right\} \right) + \mathcal{P} \left(\left\{ \sum_{i=1}^n \mathbb{E}(\rho^2(\xi_i, \xi_i^0)) \geq \frac{t^2}{4\Theta^2} \right\} \right) \\ &= I_1(t) + I_2(t). \end{aligned}$$

Notemos ahora que, como cada $\xi_i(\omega)$ toma valores en $B_\rho(\xi_i^0, \frac{R}{2n})$, entonces $0 \leq \rho^2(\xi_i(\omega), \xi_i^0) \leq \frac{R^2}{4n^2}$, para todo $\omega \in \Omega$. Por consiguiente $0 \leq \mathbb{E}(\rho^2(\xi_i, \xi_i^0)) = \int_\Omega \rho^2(\xi_i(\omega), \xi_i^0) d\mathcal{P}(\omega) \leq \frac{R^2}{4n^2}$.

Entonces, $0 \leq \sum_{i=1}^n \mathbb{E}(\rho^2(\xi_i, \xi_i^0)) \leq \frac{R^2}{4n}$, de manera que, $\frac{t^2}{4\Theta^2} > \frac{R^2}{4n}$ si $t > \frac{\Theta}{\sqrt{n}} R$, y el segundo término $I_2(t)$ es nulo. Estimemos $I_1(t)$. Sea $Z_1(\omega) = \sum_{i=1}^n [\rho^2(\xi_i(\omega), \xi_i^0) - \mathbb{E}(\rho^2(\xi_i, \xi_i^0))]$.

Podemos calcular $\Psi_{Z_1}(\lambda)$ en términos de las $\Psi_{\rho^2(\xi_i, \xi_i^0) - \mathbb{E}(\rho^2(\xi_i, \xi_i^0))}$. En efecto, usando la

independencia de las variables aleatorias $\xi_i(\omega)$, tenemos

$$\begin{aligned}
\Psi_{Z_1}(\lambda) &= \log \mathbb{E}(e^{\lambda Z_1}) \\
&= \log \mathbb{E} \left(e^{\lambda \sum_{i=1}^n [\rho^2(\xi_i, \xi_i^0) - \mathbb{E}(\rho^2(\xi_i, \xi_i^0))]} \right) \\
&= \log \mathbb{E} \left(\prod_{i=1}^n e^{\lambda [\rho^2(\xi_i, \xi_i^0) - \mathbb{E}(\rho^2(\xi_i, \xi_i^0))]} \right) \\
&= \log \prod_{i=1}^n \mathbb{E} \left(e^{\lambda [\rho^2(\xi_i, \xi_i^0) - \mathbb{E}(\rho^2(\xi_i, \xi_i^0))]} \right) \\
&= \sum_{i=1}^n \log \mathbb{E} \left(e^{\lambda [\rho^2(\xi_i, \xi_i^0) - \mathbb{E}(\rho^2(\xi_i, \xi_i^0))]} \right) \\
&= \sum_{i=1}^n \Psi_{\rho^2(\xi_i, \xi_i^0) - \mathbb{E}(\rho^2(\xi_i, \xi_i^0))}(\lambda).
\end{aligned}$$

Como las variables aleatorias $\rho^2(\xi_i, \xi_i^0) - \mathbb{E}(\rho^2(\xi_i, \xi_i^0))$ toman valores en $(-\frac{R^2}{4n^2}, \frac{R^2}{4n^2})$ podemos aplicar el Lema de Hoeffding, Teorema 12.3.2, para obtener que $\rho^2(\xi_i, \xi_i^0) - \mathbb{E}(\rho^2(\xi_i, \xi_i^0))$ es subgaussiana con factor de varianza $\frac{1}{4}(2\frac{R^2}{4n^2})^2 = \frac{R^4}{16n^4}$. En otras palabras

$$\Psi_{\rho^2(\xi_i, \xi_i^0) - \mathbb{E}(\rho^2(\xi_i, \xi_i^0))}(\lambda) \leq \frac{\lambda^2 R^4}{2 \cdot 16n^4} = \frac{\lambda^2 R^4}{32n^4}.$$

Entonces $\Psi_{Z_1}(\lambda) \leq \frac{\lambda^2 R^4}{32n^3} = \frac{\lambda^2 R^4}{2 \cdot 16n^3}$. Entonces Z_1 es subgaussiana con factor de varianza $\frac{R^4}{16n^3}$. Por consiguiente, por la Proposición 12.3.3, tenemos que

$$\begin{aligned}
I_1(t) &= \mathcal{P} \left(\left\{ Z_1(\omega) \geq \frac{t^2}{4\Theta^2} \right\} \right) \\
&\leq \exp \left(-\frac{1}{2} \frac{t^4}{16\Theta^4} \frac{16n^3}{R^4} \right) \\
&= e^{-\frac{1}{2} \frac{n^3}{\Theta^4 R^4} t^4}.
\end{aligned}$$

Resumiendo, para $t > \frac{\Theta}{\sqrt{n}} R$, tenemos que

$$I(t) = I_1(t) + I_2(t) \leq e^{-\frac{1}{2} \frac{n^3}{\Theta^4 R^4} t^4}.$$

Podemos proceder de manera similar con $II(t)$ usando ahora la independencia de las variables aleatorias $a_i(\omega)$. Notemos en primer lugar que

$$\begin{aligned}
II(t) &= \mathcal{P} \left(\left\{ \sum_{i=1}^n |a_i(\omega) - a_i^0|^2 \geq \frac{t^2}{2\Theta^2} \right\} \right) \\
&= \mathcal{P} \left(\left\{ \sum_{i=1}^n [|a_i(\omega) - a_i^0|^2 - \mathbb{E}(|a_i - a_i^0|^2)] + \sum_{i=1}^n \mathbb{E}(|a_i - a_i^0|^2) \geq \frac{t^2}{2\Theta^2} \right\} \right) \\
&\leq \mathcal{P} \left(\left\{ \sum_{i=1}^n [|a_i(\omega) - a_i^0|^2 - \mathbb{E}(|a_i - a_i^0|^2)] \geq \frac{t^2}{4\Theta^2} \right\} \right) \\
&\quad + \mathcal{P} \left(\left\{ \sum_{i=1}^n \mathbb{E}(|a_i - a_i^0|^2) \geq \frac{t^2}{4\Theta^2} \right\} \right) \\
&= II_1(t) + II_2(t).
\end{aligned}$$

Como para cada $\omega \in \Omega$ tenemos que $|a_i(\omega) - a_i^0| \leq m = \min_{i=1, \dots, n} a_i^0$, entonces $0 \leq \mathbb{E}(|a_i - a_i^0|^2) \leq m^2$. De aquí que $0 \leq \sum_{i=1}^n \mathbb{E}(|a_i - a_i^0|^2) \leq nm^2$. Por consiguiente si $t > 2m\sqrt{n}\Theta$, $II_2(t) = 0$ porque el conjunto $\left\{ \omega \in \Omega : \sum_{i=1}^n \mathbb{E}(|a_i - a_i^0|^2) > \frac{t^2}{4\Theta^2} \right\}$ es vacío. Estimemos $II_1(t)$ usando Cramér-Chernoff y Hoeffding. Sea $Z_2(\omega) = \sum_{i=1}^n [|a_i(\omega) - a_i^0|^2 - \mathbb{E}(|a_i - a_i^0|^2)]$, entonces

$$\begin{aligned}
\Psi_{Z_2}(\lambda) &= \log \mathbb{E}(e^{\lambda Z_2}) \\
&= \log \mathbb{E} \left(e^{\lambda \sum_{i=1}^n [|a_i(\omega) - a_i^0|^2 - \mathbb{E}(|a_i - a_i^0|^2)]} \right) \\
&= \log \mathbb{E} \left(\prod_{i=1}^n e^{\lambda [|a_i(\omega) - a_i^0|^2 - \mathbb{E}(|a_i - a_i^0|^2)]} \right) \\
&= \log \prod_{i=1}^n \mathbb{E} \left(e^{\lambda [|a_i(\omega) - a_i^0|^2 - \mathbb{E}(|a_i - a_i^0|^2)]} \right) \\
&= \sum_{i=1}^n \Psi_{|a_i(\omega) - a_i^0|^2 - \mathbb{E}(|a_i - a_i^0|^2)}(\lambda).
\end{aligned}$$

Por el Teorema 12.3.2, puesto que las variables aleatorias $|a_i - a_i^0|^2 - \mathbb{E}(|a_i - a_i^0|^2)$ toman valores en intervalos de longitud acotada por $2m^2$, tenemos que son subgaussianas con factor de varianza $\frac{1}{4}(2m^2)^2 = m^4$. Entonces $\Psi_{Z_2}(\lambda) \leq n\frac{\lambda}{2}m^4$, es decir Z_2 es subgaussiana con factor de varianza nm^4 . Aplicando de nuevo la Proposición 12.3.3, tenemos que,

$$II_1(t) = \mathcal{P} \left(\left\{ Z_2(\omega) > \frac{t^2}{4\Theta^2} \right\} \right)$$

$$\begin{aligned} &\leq \exp\left(-\frac{1}{2} \frac{t^4}{16\Theta^4 nm^4}\right) \\ &= e^{-\frac{1}{32} \frac{t^4}{nm^4\Theta^4}}. \end{aligned}$$

Resumiendo las estimaciones precedentes, para $t > 2m\sqrt{n}\Theta$, tenemos

$$\begin{aligned} \mathcal{P}(\{\mathfrak{d}_{GL}^{KR}(\mathcal{G}_\omega, \mathcal{G}_0) > t\}) &\leq I(t) + II(t) \\ &= e^{-\frac{1}{2} \frac{n^3}{\Theta^4 R^4} t^4} + e^{-\frac{1}{32} \frac{1}{nm^4\Theta^4} t^4}. \end{aligned}$$

□

COROLARIO 12.4.2. *En la situación general del Teorema 12.4.1, tenemos que*

a) *si $2nm = R$, o $m = \frac{R}{2n}$, entonces*

$$\mathcal{P}(\{\omega \in \Omega : \mathfrak{d}_{GL}^{KR}(\mathcal{G}_\omega, \mathcal{G}_0) > t\}) \leq 2e^{-\frac{1}{2} \frac{n^3}{\Theta^4 R^4} t^4},$$

para $t > \Theta \frac{R}{\sqrt{n}}$;

b) *si además $R = 1$, entonces $m = \frac{1}{2n}$, $\Theta = 8 + 4\sqrt{n}\gamma$ y también*

$$\mathcal{P}(\{\omega \in \Omega : \mathfrak{d}_{GL}^{KR}(\mathcal{G}_\omega, \mathcal{G}_0) > t\}) \leq 2e^{-\frac{1}{2} \frac{n^3}{\Theta^4} t^4} = 2e^{-\frac{1}{2} \frac{n^3}{(8+4\sqrt{n}\gamma)^4} t^4},$$

para $t > \frac{8+4\sqrt{n}\gamma}{\sqrt{n}}$, donde $\gamma = \text{diam}_\rho\{\xi_1^0, \dots, \xi_n^0\}$;

c) *cuando además, como en los espacios Euclídeos \mathbb{R}^k , los puntos ξ_i^0 pueden elegirse 1-dispersos con la propiedad que $\gamma \simeq n^{\frac{1}{k}}$, tenemos para $t > \frac{8}{\sqrt{n}} + 4n^{\frac{1}{k}}$, y n suficientemente grande vale la desigualdad*

$$\mathcal{P}(\{\omega \in \Omega : \mathfrak{d}_{GL}^{KR}(\mathcal{G}_\omega, \mathcal{G}_0) > t\}) \leq 2e^{-\frac{1}{16} n^{1-\frac{4}{k}} t^4}.$$

DEMOSTRACIÓN. a) Tomando $2nm = R$, los dos términos de la derecha en la desigualdad de la tesis del teorema anterior, tenemos que $\frac{1}{2} \frac{n^3}{\Theta^2 R^4} t^4 = \frac{1}{32} \frac{t^4}{n\Theta^2 m^4}$ para todo t y así obtenemos a).

b) Si además $R = 1$, tenemos que $\Theta = 4 \max\{2, 2 + \sqrt{n}\gamma\} = 8 + 4\sqrt{n}\gamma$. También $\Theta \max\left\{\frac{R}{\sqrt{n}}, 2m\sqrt{n}\right\} = (8 + 4\sqrt{n}\gamma) = \frac{8}{\sqrt{n}} + 4\gamma$. Por consiguiente, de a), si $t > \frac{8}{\sqrt{n}} + 4\gamma$, entonces

$$\mathcal{P}(\{\omega \in \Omega : \mathfrak{d}_{GL}^{KR}(\mathcal{G}_\omega, \mathcal{G}_0) > t\}) \leq 2e^{-\frac{1}{2} \frac{n^3}{(8+4\sqrt{n}\gamma)^4} t^4}.$$

c) Notemos primero que, aunque en algunos casos el diámetro de un conjunto $R = 1$ disperso con n puntos es del orden de n , en general puede ser mucho más chico. Así en \mathbb{R}^k se pueden tomar n puntos 1 dispersos de manera que el diámetro del conjunto de esos n puntos sea del orden de $n^{\frac{1}{k}}$. Si ese es el caso, para n suficientemente grande tenemos que el comportamiento como función de n del exponente en la desigualdad en b) será del orden de

$$-\frac{n^3 t^4}{\left(n^{\frac{1}{2}} n^{\frac{1}{k}}\right)^4} = -n^{1-\frac{4}{k}} t^4$$

lo que prueba c). □

Observamos que, como $d_{GL}^{KR}(\mathcal{G}_\omega, \mathcal{G}_0)$ resulta acotado por arriba, entonces el conjunto $\{\omega \in \Omega : d_{GL}^{KR}(\mathcal{G}_\omega, \mathcal{G}_0) > t\}$ es vacío para t suficientemente grande y la estimación del Teorema 12.4.1 es trivial. No obstante, los resultados precedentes dan una estimación cuantitativa para valores de t que a priori no necesariamente supone esa cota cuyo valor óptimo no es fácil de estimar por la propia definición de d_{GL}^{KR} .

12.5. Concentración espectral de grafos de afinidades con atributos aleatorios

Consideramos ahora la situación descrita en general en la Sección 10.2, es decir, tenemos variables aleatorias valuadas en grafos de afinidades con atributos. Dado $(\Omega, \mathcal{F}, \mathcal{P})$ un espacio de probabilidad, para $\omega \in \Omega$ tenemos un grafo $\mathcal{G}_\omega = (\mathcal{V}, \mathcal{E}, \bar{a}(\omega), W(\omega))$. Para que \mathcal{G}_ω pueda ser visto como un grafo ponderado no dirigido los $w_{ij}(\omega)$ no pueden ser negativos, mientras que los $a_i(\omega)$ deben ser positivos. Si, $w_{ij} \geq 0$, $a_i > 0$ y todas son subgaussianas, es posible estimar la concentración del Laplaciano asociado al grafo \mathcal{G}_ω alrededor del grafo medio \mathcal{G}_0 . Precisamos el contexto en el próximo resultado algunas de cuyas afirmaciones están en [AAT23].

TEOREMA 12.5.1. *Sea $(\Omega, \mathcal{F}, \mathcal{P})$ un espacio de probabilidad. Sea $\mathcal{V} = \{1, \dots, n\}$. Sean $w_{ij} : \Omega \rightarrow \mathbb{R}_0^+$, $\frac{n(n-1)}{2}$ -variables aleatorias independientes para $1 \leq i < j \leq n$, que toman valores en el intervalo $[0, \alpha]$, $\alpha > 0$. Sean $a_i : \Omega \rightarrow \mathbb{R}^+$ n -variables aleatorias independientes que toman valores en el intervalo $[\beta, \gamma]$ con $0 < \beta < \gamma < \infty$. Suponemos también que las variables aleatorias $a_i(\omega)$ y las $w_{ij}(\omega)$ son independientes. Entonces, si Δ_ω denota el Laplaciano asociado al grafo $\mathcal{G}_\omega = (\mathcal{V}, \mathcal{E}, \bar{a}(\omega), W(\omega))$, Δ denota el Laplaciano*

asociado al grafo $\mathcal{G}_0 = (\mathcal{V}, \mathcal{E}, \mathbb{E}(\bar{a}), \mathbb{E}(W))$, $\bar{f} = (f(1), \dots, f(n))$, $\bar{f}_i = (f(i), \dots, f(i))$ y $\|\cdot\|$ denota la norma Euclídea en \mathbb{R}^n , tenemos que

(1) la desigualdad

$$\mathcal{P}(\{\Delta_\omega f(i) - \Delta f(i) > t\}) \leq e^{-\frac{2t^2}{\alpha^2[5\beta^2 + \gamma^2 - 2\gamma\beta]} \frac{1}{\|\bar{f} - \bar{f}_i\|^2}}$$

vale para todo $i \in \mathcal{V}$ y todo $t > 0$;

(2) la desigualdad

$$\mathcal{P}(\{\Delta_\omega f(i) - \Delta f(i) < -t\}) \leq e^{-\frac{2t^2}{\alpha^2[5\beta^2 + \gamma^2 - 2\gamma\beta]} \frac{1}{\|\bar{f} - \bar{f}_i\|^2}}$$

vale para todo $i \in \mathcal{V}$ y todo $t > 0$;

(3) la desigualdad

$$\mathcal{P}(\{\|\Delta_\omega f - \Delta f\| > t\}) \leq ne^{-\frac{2t^2}{n^2\alpha^2[5\beta^2 + \gamma^2 - 2\gamma\beta]} \frac{1}{(\max f - \min f)^2}}$$

vale para todo $t > 0$;

(4) si ϕ es una autofunción de Δ con autovalor λ y $\|\phi\| = 1$ entonces la desigualdad

$$\mathcal{P}(\{\|\Delta_\omega \phi - \lambda \phi\| > t\}) \leq ne^{-\frac{2t^2}{n^2\alpha^2[5\beta^2 + \gamma^2 - 2\gamma\beta]} \frac{1}{(\max \phi - \min \phi)^2}}$$

vale para todo $t > 0$.

DEMOSTRACIÓN. (1) Sean f una función definida en los vértices e i en \mathcal{V} fijos. Entonces $\Delta f(i) = \frac{1}{\mathbb{E}(a_i)} \sum_{j=1}^n \mathbb{E}(w_{ij})(f(j) - f(i))$ y $\Delta_\omega f(i) = \frac{1}{a_i(\omega)} \sum_{j=1}^n w_{ij}(\omega)(f(j) - f(i))$. Notar que

(12.2)

$$\begin{aligned} & \mathcal{P}(\{\omega \in \Omega : \Delta_\omega f(i) - \Delta f(i) > t\}) = \\ & = \mathcal{P}\left(\left\{\omega \in \Omega : \sum_{j=1}^n w_{ij}(\omega)(f(j) - f(i)) - \frac{a_i(\omega)}{\mathbb{E}(a_i)} \sum_{j=1}^n \mathbb{E}(w_{ij})(f(j) - f(i)) > ta_i(\omega)\right\}\right) \\ & \leq \mathcal{P}\left(\left\{\omega \in \Omega : \sum_{j=1}^n \left(w_{ij}(\omega) - \mathbb{E}(w_{ij}) \frac{a_i(\omega)}{\mathbb{E}(a_i)}\right) (f(j) - f(i)) > t\beta\right\}\right). \end{aligned}$$

Por otra parte

$$\mathbb{E}\left(w_{ij} - \mathbb{E}(w_{ij}) \frac{a_i}{\mathbb{E}(a_i)}\right) (f(j) - f(i)) = \left[\mathbb{E}(w_{ij}) - \mathbb{E}(w_{ij}) \frac{\mathbb{E}(a_i)}{\mathbb{E}(a_i)}\right] (f(j) - f(i)) = 0.$$

Probemos ahora que la variable aleatoria centrada

$$Z(\omega) = \sum_{j=1}^n \left(w_{ij}(\omega) - \mathbb{E}(w_{ij}) \frac{a_i(\omega)}{\mathbb{E}(a_i)} \right) (f(j) - f(i))$$

es subgaussiana con factor de varianza $\nu = \alpha^2 \left[1 + \frac{(\gamma - \beta)^2}{4\beta^2} \right] \|\bar{f} - \bar{f}_i\|^2$. Calculemos entonces

$$\begin{aligned} \Psi_Z(\lambda) &= \log \mathbb{E} (e^{\lambda Z}) \\ &= \log \mathbb{E} \left(\prod_{j=1}^n e^{\lambda(f(j)-f(i)) \left(w_{ij}(\omega) - \frac{\mathbb{E}(w_{ij})}{\mathbb{E}(a_i)} a_i(\omega) \right)} \right) \\ &= \log \mathbb{E} \left(\prod_{j=1}^n \left[e^{\lambda(f(j)-f(i))w_{ij}(\omega)} e^{-\lambda(f(j)-f(i)) \frac{\mathbb{E}(w_{ij})}{\mathbb{E}(a_i)} a_i(\omega)} \right] \right) \\ &= \log \left[\prod_{j=1}^n \mathbb{E} \left(e^{\lambda(f(j)-f(i))w_{ij}(\omega)} \right) \mathbb{E} \left(e^{-\lambda(f(j)-f(i)) \frac{\mathbb{E}(w_{ij})}{\mathbb{E}(a_i)} a_i(\omega)} \right) \right] \\ &= \sum_{j=1}^n \Psi_{(f(j)-f(i))w_{ij}}(\lambda) + \sum_{j=1}^n \Psi_{(f(i)-f(j)) \frac{\mathbb{E}(w_{ij})}{\mathbb{E}(a_i)} a_i}(\lambda). \end{aligned}$$

Estimemos ahora, por Hoeffding las funciones $\Psi_{(f(j)-f(i))w_{ij}}(\lambda)$ y $\Psi_{(f(i)-f(j)) \frac{\mathbb{E}(w_{ij})}{\mathbb{E}(a_i)} a_i}(\lambda)$. Recordemos para ello que las w_{ij} toman valores en $[0, \alpha]$ y las a_i toman valores en $[\beta, \gamma]$. Entonces $(f(j) - f(i))w_{ij}$ toman valores en $[-|f(j) - f(i)|\alpha, |f(j) - f(i)|\alpha]$ cuya longitud es $2|f(j) - f(i)|\alpha$. Así

$$\Psi_{(f(j)-f(i))w_{ij}}(\lambda) \leq \frac{\lambda^2}{2} |f(j) - f(i)|^2 \alpha^2.$$

Análogamente para $\Psi_{(f(i)-f(j)) \frac{\mathbb{E}(w_{ij})}{\mathbb{E}(a_i)} a_i}(\lambda)$, tenemos que $(f(i) - f(j)) \frac{\mathbb{E}(w_{ij})}{\mathbb{E}(a_i)} a_i$ está en el intervalo $\left[-\gamma|f(i) - f(j)| \frac{\mathbb{E}(w_{ij})}{\mathbb{E}(a_i)}, -\beta|f(i) - f(j)| \frac{\mathbb{E}(w_{ij})}{\mathbb{E}(a_i)} \right]$ o bien $(f(i) - f(j)) \frac{\mathbb{E}(w_{ij})}{\mathbb{E}(a_i)} a_i$ está en $\left[\beta|f(i) - f(j)| \frac{\mathbb{E}(w_{ij})}{\mathbb{E}(a_i)}, \gamma|f(i) - f(j)| \frac{\mathbb{E}(w_{ij})}{\mathbb{E}(a_i)} \right]$. En cualquier caso el factor de varianza será $\frac{1}{4}(\gamma - \beta)^2 |f(i) - f(j)|^2 \left(\frac{\mathbb{E}(w_{ij})}{\mathbb{E}(a_i)} \right)^2 \leq \frac{1}{4} \frac{\alpha^2}{\beta^2} (\gamma - \beta)^2 |f(i) - f(j)|^2$. Por consiguiente

$$\Psi_{(f(i)-f(j)) \frac{\mathbb{E}(w_{ij})}{\mathbb{E}(a_i)} a_i}(\lambda) \leq \frac{\lambda^2 \alpha^2}{8 \beta^2} (\gamma - \beta)^2 |f(i) - f(j)|^2.$$

Resumiendo las estimaciones anteriores, tenemos que

$$\Psi_Z(\lambda) \leq \sum_{j=1}^n \frac{\lambda^2}{2} |f(j) - f(i)|^2 \alpha^2 + \sum_{j=1}^n \frac{\lambda^2 \alpha^2}{8 \beta^2} (\gamma - \beta)^2 |f(i) - f(j)|^2$$

$$\begin{aligned}
&= \left[\frac{\lambda^2}{2} \alpha^2 + \frac{\lambda^2}{8} (\gamma - \beta)^2 \frac{\alpha^2}{\beta^2} \right] \sum_{j=1}^n |f(j) - f(i)|^2 \\
&= \frac{\lambda^2}{2} \alpha^2 \left[1 + \frac{(\gamma - \beta)^2}{4\beta^2} \right] \|\bar{f} - \bar{f}_i\|^2.
\end{aligned}$$

Por lo tanto Z es subgaussiana con factor de varianza $\nu = \alpha^2 \left[1 + \frac{(\gamma - \beta)^2}{4\beta^2} \right] \|\bar{f} - \bar{f}_i\|^2$. Por la Proposición 12.3.3 y por (12.2) tenemos que

$$\begin{aligned}
\mathcal{P}(\{\Delta_\omega f(i) - \Delta f(i) > t\}) &\leq \mathcal{P}(\{Z > t\beta\}) \\
&\leq e^{-\frac{t^2 \beta^2}{2} \frac{1}{\nu}} \\
&= e^{-\frac{t^2 \beta^2}{2} \frac{1}{\alpha^2 \left[1 + \frac{(\gamma - \beta)^2}{4\beta^2} \right] \|\bar{f} - \bar{f}_i\|^2}} \\
&= e^{-\frac{2t^2}{\alpha^2 [5\beta^2 + \gamma^2 - 2\gamma\beta]} \frac{1}{\|\bar{f} - \bar{f}_i\|^2}}.
\end{aligned}$$

(2) Sigue de (1) tomando $-f$ en lugar de f . Veamos (3)

$$\begin{aligned}
\mathcal{P}(\{\|\Delta_\omega f - \Delta f\| > t\}) &= \mathcal{P}\left(\left\{\sum_{i=1}^n |\Delta_\omega f(i) - \Delta f(i)|^2 > t^2\right\}\right) \\
&\leq \mathcal{P}\left(\bigcup_{i=1}^n \left\{|\Delta_\omega f(i) - \Delta f(i)|^2 > \frac{t^2}{n}\right\}\right) \\
&= \mathcal{P}\left(\bigcup_{i=1}^n \left\{|\Delta_\omega f(i) - \Delta f(i)| > \frac{t}{\sqrt{n}}\right\}\right) \\
&\leq \sum_{i=1}^n \mathcal{P}\left(\left\{|\Delta_\omega f(i) - \Delta f(i)| > \frac{t}{\sqrt{n}}\right\}\right) \\
&\leq \sum_{i=1}^n e^{-\frac{2t^2}{n\alpha^2(5\beta^2 + \gamma^2 - 2\gamma\beta)} \frac{1}{\|\bar{f} - \bar{f}_i\|^2}} \\
&\leq ne^{-\frac{2t^2}{n^2\alpha^2(5\beta^2 + \gamma^2 - 2\gamma\beta)} \frac{1}{(\max f - \min f)^2}}.
\end{aligned}$$

Para (4) usamos (3) sustituyendo f por ϕ y $\Delta\phi$ por $\lambda\phi$. □

Observamos que en las tres primeras desigualdades en (1), (2) y (3) del Teorema 12.5.1 el exponente contiene factores de las formas $\frac{1}{\|\bar{f} - \bar{f}_i\|^2}$ y $\frac{1}{(\max f - \min f)^2}$. Cuando estos números son grandes, por que f es una señal de baja amplitud, tendremos estimaciones mejores para las desviaciones puntuales y en norma de $\Delta_\omega f$ con respecto a Δf .

Observamos que el ítem (4) del teorema precedente estima que la probabilidad de grandes desviaciones espectrales del Laplaciano aleatorio alrededor del Laplaciano de las medias. Pero la desigualdad obtenida no es apta para ser usada para estimar grandes desviaciones de las métricas difusivas alrededor de la métrica difusiva asociada al grafo medio. La razón es que en las métricas difusivas los autovalores y las correspondientes autofunciones juegan papeles bien diferenciados. En (4) en cambio sólo vemos desviación con respecto al producto $\lambda\phi$. Cuando fuera posible obtener modelos cuyas autofunciones están fijas y sólo los autovalores son aleatorios, todavía es simple obtener estimaciones de concentración para las distancias difusivas. La idea geométrica corresponde con perturbaciones en los semidiámetros de elipsoides con direcciones principales fijas, ver Figura 12.5.1. Consideremos brevemente esta situación particular. Sea, como antes $\mathcal{V} = \{1, 2, \dots, n\} \in \mathbb{N}$. Sean $\phi_0, \phi_1, \dots, \phi_{n-1}$ ortonormales con el producto interno usual de \mathbb{R}^n y con $\phi_0(i) = \phi_0(j)$, para todo $i, j \in \mathcal{V}$. Consideramos que los $n - 1$ autovalores no triviales λ_j con $j = 1, \dots, n - 1$, son negativos aleatorios e independientes pero que las autofunciones están fijas y son las correspondientes $\phi_1, \dots, \phi_{n-1}$.

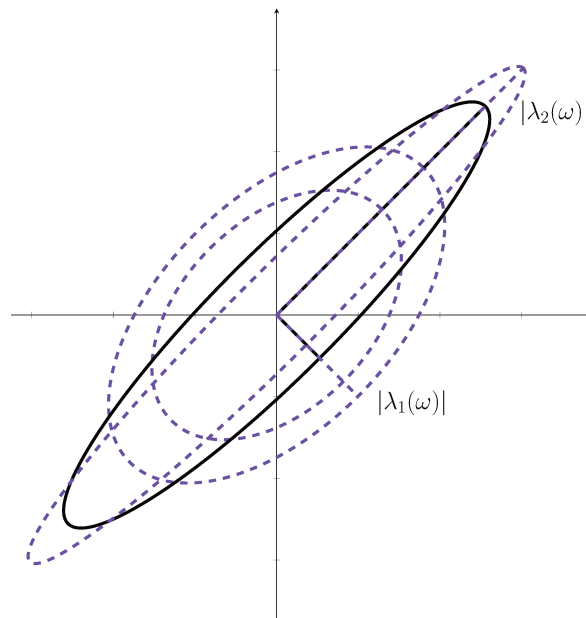


FIGURA 12.5.1. Autovectores fijos, autovalores aleatorios.

Sea μ_j la distribución de $-\lambda_j(\omega) = |\lambda_j(\omega)|$, es decir que $\mu_j(B) = \mathcal{P}(\{-\lambda_j \in B\})$ con $B \in \mathcal{B}(\mathbb{R})$. Notemos que las μ_j tienen soportes en \mathbb{R}^+ , $j = 1, \dots, n - 1$.

LEMA 12.5.2. *Para cada $t > 0$ y para cada $i \in \mathcal{V}$ la variable aleatoria $e^{2t\lambda_i(\omega)}$ tiene esperanza finita y*

$$\mathbb{E}(e^{2t\lambda_i}) = \mathcal{L}\mu_i(2t),$$

donde $\mathcal{L}\mu_i(s) = \int_{\mathbb{R}} e^{-sx} d\mu_i(x)$, es la transformada de Laplace de la distribución μ_i de $-\lambda_i$.

DEMOSTRACIÓN. Para $t > 0$, como cada λ_i toma sólo valores negativos, las variables aleatorias $e^{2t\lambda_i(\omega)}$ toman valores en el intervalo $[0, 1]$, por consiguiente tienen todos los momentos finitos. Si $\bar{\mu}_i(B) = \mu_i(-B)$ con $B \in \mathcal{B}(\mathbb{R})$, entonces $\bar{\mu}_i$ es la distribución de λ_i y

$$\begin{aligned} \mathbb{E}(e^{2t\lambda_i}) &= \int_{\mathbb{R}} e^{2tx} d\bar{\mu}_i(x) \\ &= \int_{\mathbb{R}} e^{-2tx} d\mu_i(x) \\ &= \mathcal{L}\mu_i(2t). \end{aligned}$$

□

COROLARIO 12.5.3. *Si $d_t(\omega)$ denota la métrica difusiva aleatoria asociada a los autovalores $\{\lambda_i(\omega); i = 1, \dots, n-1\}$ y las autofunciones fijas $\{\phi_i; i = 1, \dots, n-1\}$ dadas por $d_t^2(\omega)(i, j) = \sum_{l=1}^{n-1} e^{2t\lambda_l(\omega)} |\phi_l(i) - \phi_l(j)|^2$, entonces*

$$\mathbb{E}(d_t^2(\omega)(i, j)) = \sum_{l=1}^{n-1} \mathcal{L}\mu_l(2t) |\phi_l(i) - \phi_l(j)|^2.$$

DEMOSTRACIÓN. Basta aplicar el lema anterior. □

Estamos en condiciones de estudiar la concentración de $d_t^2(\omega)$ alrededor de $\mathbb{E}(d_t^2)$ cuando los autovalores son variables aleatorias independientes.

TEOREMA 12.5.4. *Sea $\mathcal{V} = \{1, 2, \dots, n\}$, sean $\phi_1, \dots, \phi_{n-1}$, $n-1$ vectores ortonormales de \mathbb{R}^n todos ortogonales al vector $(1, \dots, 1) \in \mathbb{R}^n$. Sean $\{\lambda_i(\omega); i = 1, \dots, n-1\}$ $n-1$ variables aleatorias independientes con valores negativos y sea μ_i la distribución de $-\lambda_i$ con $i = 1, \dots, n-1$. Entonce la desigualdad*

$$\mathcal{P}(\{\omega \in \Omega : d_t^2(\omega)(i, j) - \mathbb{E}(d_t^2(i, j)) > s\}) \leq e^{-\frac{s^2}{2} \frac{1}{\sum_{l=1}^{n-1} |\phi_l(i) - \phi_l(j)|^4}} \leq e^{-\frac{s^2}{32(n-1)}},$$

vale para todo $s > 0$.

DEMOSTRACIÓN. Observamos que, como $t > 0$ y $\lambda_i(\omega) < 0$ para $\omega \in \Omega$, entonces $e^{2t\lambda_i(\omega)} - \mathcal{L}\mu_i(2t) = e^{2t\lambda_i(\omega)} - \mathbb{E}(e^{2t\lambda_i})$ toma valores en el intervalo $[-1, 1]$ ya que $0 \leq \mathbb{E}(e^{2t\lambda_i}) = \int_{\Omega} e^{2t\lambda_i(\omega)} d\mathcal{P}(\omega) \leq 1$. Por el Lema de Hoeffding, Teorema 12.3.2, la variable aleatoria $e^{2t\lambda_i(\omega)} - \mathcal{L}\mu_i(2t)$ es subgaussiana con factor de varianza $\frac{(1-(-1))^2}{4} = 1$. En otras palabras

$$\Psi_{e^{2t\lambda_i} - \mathcal{L}\mu_i(2t)}(\xi) \leq \frac{\xi^2}{2}.$$

Estimemos, como antes, $\Psi_{Z(i,j)}$ para

$$Z(\omega)(i, j) = d_t^2(\omega)(i, j) - \mathbb{E}(d_t^2(i, j)) = \sum_{l=1}^{n-1} [e^{2t\lambda_l(\omega)} - \mathcal{L}\mu_l(2t)] |\phi_l(i) - \phi_l(j)|^2,$$

En efecto, usando la independencia de las variables aleatorias $\lambda_l(\omega)$,

$$\begin{aligned} \Psi_{Z(\omega)(i,j)}(\xi) &= \log \mathbb{E} \left(e^{\xi Z(\omega)(i,j)} \right) \\ &= \log \mathbb{E} \left(e^{\xi \sum_{l=1}^{n-1} [e^{2t\lambda_l} - \mathcal{L}\mu_l(2t)] |\phi_l(i) - \phi_l(j)|^2} \right) \\ &= \log \mathbb{E} \left(\prod_{l=1}^{n-1} e^{\xi [e^{2t\lambda_l} - \mathcal{L}\mu_l(2t)] |\phi_l(i) - \phi_l(j)|^2} \right) \\ &= \log \prod_{j=1}^{n-1} \mathbb{E} \left(e^{\xi [e^{2t\lambda_l} - \mathcal{L}\mu_l(2t)] |\phi_l(i) - \phi_l(j)|^2} \right) \\ &= \sum_{l=1}^{n-1} \Psi_{e^{2t\lambda_l} - \mathcal{L}\mu_l(2t)}(\xi |\phi_l(i) - \phi_l(j)|^2) \\ &\leq \sum_{l=1}^{n-1} \frac{\xi^2 |\phi_l(i) - \phi_l(j)|^4}{2} \\ &= \frac{\xi^2}{2} \left(\sum_{l=1}^{n-1} |\phi_l(i) - \phi_l(j)|^4 \right). \end{aligned}$$

En otros términos, la variable aleatoria $Z(\omega)(i, j)$ resulta subgaussiana con factor de varianza $\sum_{l=1}^{n-1} |\phi_l(i) - \phi_l(j)|^4$. Por la Proposición 12.3.3 obtenemos que

$$\mathcal{P}(\{Z(\omega)(i, j) > s\}) \leq e^{-\frac{s^2}{2 \sum_{l=1}^{n-1} |\phi_l(i) - \phi_l(j)|^4}},$$

para todo $s > 0$. Finalmente, puesto que los vectores ϕ_l tienen norma igual a uno, cada $|\phi_l(i)|$ es menor o igual que uno. Entonces $\sum_{l=1}^{n-1} |\phi_l(i) - \phi_l(j)|^4 \leq (n-1)16$. \square

Conclusiones

Conclusiones teóricas centrales

- (1) Algunas clases de espacios métricos probabilísticos se pueden metrizar a través del método de Gromov-Lipschitz en las métricas y el método de Kantorovich-Rubinstein-Wasserstein en las medidas de probabilidad.
- (2) Con la estructura del punto (1) se pueden considerar espacios métricos de probabilidad como objetos aleatorios.
- (3) Cuando los espacios básicos son grafos métricos o de afinidades las variables aleatorias adquieren forma computable y los momentos adquieren sentido.
- (4) Cuando los atributos y las afinidades están adecuadamente distribuidas se obtienen estimaciones para la concentración y la desviación alrededor de “valores medios” que garantizan estabilidad del modelo frente a muestreos aleatorios subgaussianos de los datos que lo construyen.
- (5) Algunas afinidades producen métricas. Otras no, pero lo mismo generan topologías por sistemas de entornos cuyos umbralados proveen dinámicas topológicas reveladoras de algunas propiedades de estructuras de los datos.

Conclusiones para la aplicación a SUBE en AMBA

- (6) Las diferencias de concavidades (derivadas segundas) de las curvas de dinámica topológica, medida en términos de la evolución de la cantidad de componentes conexas como función del parámetro de umbralado, refleja diferenciación en la distribución geográfica de población. Esto puede proveer un parámetro útil para

diferenciar grafos ponderados.

- (7) Los resultados teóricos proveen, bajo ciertas condiciones, estabilidad para el modelo difusivo de AMBA basado en SUBE.
- (8) Los cálculos computacionales muestran diferencias espectrales de comportamiento en distintas etapas tempranas de la pandemia.
- (9) El impacto en la estructura espectral de atributos particulares en los 41 distritos de AMBA puede proveer marcadores útiles en el análisis del transporte público en relación con la propagación de pandemias.
- (10) Los aspectos teóricos y aplicados dejan abiertas muchas posibilidades de trabajo futuro.

Bibliografía

- [AAG23] Hugo Aimar, Carlos Exequiel Arias, and Ivana Gómez, *Topological evolution with increasing thresholding of undirected weighted graphs and applications*, Revista Matemática Aplicada, Computacional e Industrial (MACI), Proceedings of IX MACI **9** (2023), 291–294.
- [AAGM22] M. F. Acosta, H. Aimar, I. Gómez, and F. Morana, *Diffusive metrics induced by random affinities on graphs. an application to the transport systems related to the COVID-19 setting for Buenos aires (AMBA)*, Trends Comput. Appl. Math. **23** (2022), no. 4, 783–799. MR 4537089
- [AAT23] Carlos Exequiel Arias, Hugo Aimar, and Diego Tomassi, *Concentración del Laplaciano en grafos aleatorios*, Duodécima Jornadas Universitarias de Ciencias Exactas y Naturales, 1a. ed. Catamarca: Universidad Nacional de Catamarca. Facultad de Ciencias Exactas y Naturales (2023), 3–14.
- [ACI09] Hugo Aimar, Marilina Carena, and Bibiana Iaffei, *Discrete approximation of spaces of homogeneous type*, J. Geom. Anal. **19** (2009), no. 1, 1–18. MR 2465294
- [ACI10] ———, *Completeness of Muckenhoupt classes*, J. Math. Anal. Appl. **361** (2010), no. 2, 401–410. MR 2568704
- [AG18] Hugo Aimar and Ivana Gómez, *Affinity and distance. On the Newtonian structure of some data kernels*, Anal. Geom. Metr. Spaces **6** (2018), no. 1, 89–95. MR 3816950
- [AG22] ———, *Boundedness and concentration of random singular integrals defined by wavelet summability kernels*, J. Math. Anal. Appl. **514** (2022), no. 2, Paper No. 126315, 16. MR 4422401
- [Aim] Hugo Aimar, *Distance and Measure in Analysis and PDE*, en redacción.
- [BBI01] Dmitri Burago, Yuri Burago, and Sergei Ivanov, *A course in metric geometry*, Graduate Studies in Mathematics, vol. 33, American Mathematical Society, Providence, RI, 2001. MR 1835418
- [BBL⁺17] Michael M Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst, *Geometric deep learning: going beyond euclidean data*, IEEE Signal Processing Magazine **34** (2017), no. 4, 18–42.
- [Bil71] Patrick Billingsley, *Weak convergence of measures: Applications in probability*, Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 5, Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1971. MR 0310933

- [BLM13] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart, *Concentration inequalities*, Oxford University Press, Oxford, 2013, A nonasymptotic theory of independence, With a foreword by Michel Ledoux. MR 3185193
- [Bor44] Jorge Luis Borges, *Ficciones*, Emecé Editores, 1944.
- [Chu01] Kai Lai Chung, *A course in probability theory*, third ed., Academic Press, Inc., San Diego, CA, 2001. MR 1796326
- [CL06] Ronald R. Coifman and Stéphane Lafon, *Diffusion maps*, Appl. Comput. Harmon. Anal. **21** (2006), no. 1, 5–30. MR 2238665
- [CT06] Thomas M. Cover and Joy A. Thomas, *Elements of information theory*, second ed., Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, 2006. MR 2239987
- [Gro99] Misha Gromov, *Metric structures for Riemannian and non-Riemannian spaces*, Progress in Mathematics, vol. 152, Birkhäuser Boston, Inc., Boston, MA, 1999, Based on the 1981 French original [MR0682063 (85e:53051)], With appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael Bates. MR 1699320
- [GS02] Alison L. Gibbs and Francis Edward Su, *On choosing and bounding probability metrics*, International Statistical Review/ Revue Internationale de Statistique **70** (2002), no. 3, 419–435.
- [Gus74] Jan Gustavsson, *Metriization of quasi-metric spaces*, Math. Scand. **35** (1974), 56–60. MR 358715
- [Her96] I. N. Herstein, *Abstract algebra*, third ed., Prentice Hall, Inc., Upper Saddle River, NJ, 1996, With a preface by Barbara Cortzen and David J. Winter. MR 1375019
- [Hoe63] Wassily Hoeffding, *Probability inequalities for sums of bounded random variables*, J. Amer. Statist. Assoc. **58** (1963), 13–30. MR 144363
- [Kel62] John L. Kelley, *Topología general*, Editorial Universitaria de Buenos Aires, Buenos Aires, 1962, Translated from the English and revised by Oscar A. Varsavsky. MR 0198408
- [Mí12] Facundo Mémoli, *Some properties of Gromov-Hausdorff distances*, Discrete Comput. Geom. **48** (2012), no. 2, 416–440. MR 2946454
- [MS79] Roberto A. Macías and Carlos Segovia, *Lipschitz functions on spaces of homogeneous type*, Adv. in Math. **33** (1979), no. 3, 257–270. MR 546295
- [PC19] Gabriel Peyré and Marco Cuturi, *Computational Optimal Transport: with applications to data science*, Foundations and Trends in Machine Learning **11** (2019), no. 5-6, 355–607.
- [Tay97] J. C. Taylor, *An introduction to measure and probability*, Springer-Verlag, New York, 1997. MR 1420194
- [Vil09] Cédric Villani, *Optimal transport*, Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 338, Springer-Verlag, Berlin, 2009, Old and new. MR 2459454

-
- [WB19] Jonathan Weed and Francis Bach, *Sharp asymptotic and finite-sample rates of convergence of empirical measures in Wasserstein distance*, *Bernoulli* **25** (2019), no. 4A, 2620–2648.
MR 4003560
- [Wil88] J. H. Wilkinson, *The algebraic eigenvalue problem*, *Monographs on Numerical Analysis*, The Clarendon Press, Oxford University Press, New York, 1988, Oxford Science Publications.
MR 950175