

UNIVERSITAT DE VALÈNCIA

Facultad de Matemáticas

Departamento de Estadística e Investigación Operativa



Geoestadística en regiones heterogéneas
con distancia basada en el coste

TESIS DOCTORAL

Facundo Martín Muñoz Viera

Director: Antonio López Quílez

Burjassot, 2012

*A Carmen y Homero, a Viqui, a Luana y Nerea.
Mis antes, mi durante y mis despueses.*

The real challenge in today's multidisciplinary scientific arena is not merely to develop mathematical techniques to deal with numerical data, but also to develop the theoretical means for interpreting and integrating these data –as well as other important sources of physical knowledge– into the process of understanding.

George Christakos, 2000.

Reconocimientos

Me siento muy afortunado por tener tanto que escribir en esta sección. Aunque soy el autor material de las páginas que siguen —y de paso, el único responsable de todos los errores, inexactitudes y omisiones—, este trabajo tiene muchísimos autores intelectuales, cómplices e inductores.

Quiero empezar por mis padres biológicos (tengo más, también soy afortunado en esto), Carmen Viera y Homero Muñoz, verdaderos cabecillas de la trama, que sin tener esta Tesis en concreto como objetivo particular, sembraron en mí una curiosidad infinita y me enseñaron con el ejemplo el formidable placer que se esconde detrás del entendimiento y el descubrimiento. Mi mamá, trabajadora infatigable y perseverante irreductible, fue mi pilar fundamental, el que siempre estuvo ahí, firme y seguro. Mi papá, siempre brillante, omnisciente y audaz, me mostró el mundo, me voló la cabeza mil veces, me enseñó a desconocer las fronteras. Ellos también tuvieron sus colaboradores. Quiero mencionar especialmente a Guzmán Besada, mi papá putativo, que supo llenar todos los huecos.

En mis primeros años de estudio, en Montevideo, tuve la inmensa suerte de conocer a muchos profesores verdaderamente extraordinarios. Para empezar, el viejo y eterno Ceyser Olivera, que en el bachillerato me enseñó Matemáticas de verdad, desde el principio, paso a paso, sin dejarse nada. Desde entonces, más para bien que para mal, no supe aprender de otra manera. Con él disfruté de conocer todos los rincones de los polinomios, las cónicas y las funciones. *Al Ceyser lo que es del Ceyser.*

VIII

Ya en la Facultad de Ingeniería, la querida FING, me encontré con una tropa increíble de gigantes. Matemáticos de primera línea como Roberto Markarián, capaz de dar una clase a 300 personas sobre cualquier tema y hacerlo interesantísimo. O Jorge Groisman, el más Docente de los docentes, experto en transmutar lo ininteligible en natural. Y por supuesto, mi querido, entrañable, inmenso Gonzalo Perera. Gonzalo es de esos tipos que no pueden existir, porque lo es todo: buen matemático, músico, humanista, excelente persona, culto pero cercano, divertido, buen profesor. A Gonzalo le debo mi iniciación en la Estadística, cuando trabajé como becario en el Laboratorio de Probabilidad y Estadística que él dirigía. Pero eso es lo de menos. A todos ellos, y a algunos más, les debo fundamentalmente el haber querido ser como ellos, el saber que se puede ser así de bueno, de integral, de solidario, de comprometido. Muchos de quienes no han tenido la suerte de conocer gente así sucumben ante el mensaje incesante del *no se puede* y se lo acaban creyendo.

Pero también quiero reconocer el esfuerzo y la buena voluntad de muchos otros que no fueron tan extraordinarios, de quienes tal vez no recuerdo sus nombres, y cuya contribución anónima merece el mayor de mis respetos.

Las circunstancias de la vida acabaron por traerme a la Universitat de València, donde también tuve la oportunidad de conocer gente maravillosa. Desde mi punto de vista de estudiante de Estadística, David Conesa, Carmen Armero y Antonio López fueron tres profesores que me hechizaron. Por lo que hacían, y sobre todo por cómo lo hacían. Cuando me enteré de que formaban un grupo de investigación supe que quería trabajar con ellos, aprender de ellos.

Siempre le estaré infinitamente agradecido a Antonio por haberme abierto esa puerta, permitiéndome disfrutar del ambiente de trabajo y camaradería excepcionales en el *Grup d'Estadística espacial i temporal en Epidemiologia i Medi Ambient* (GEeitEma). A los pocos meses de empezar a trabajar tuvimos un workshop interno del grupo en Benimaurell, Alicante (Second Workshop on Bayesian Spatio-Temporal Models), que consistía en un fin de semana con la familia en un hotel precioso en lo alto de una montaña. Dedicábamos las mañanas a trabajar y las tardes a disfrutar, a jugar, a divertirnos. Si aquello no

era la felicidad, se le parecía mucho. No hay salario que pague aquellos momentos y los que vinieron después (a los que me fui acostumbrando peligrosamente). Estoy enormemente agradecido con toda aquella gente: Migue, Paloma, Gonzalo, Xavi, Susie, Asun, Javi. Ha sido una experiencia fabulosa haberlos conocido, y aprendido tanto de todos ustedes.

Por supuesto, la experiencia cotidiana tengo que agradecerse sobre todo a mis queridos compañeros de despacho. En primer lugar a *la jefa*, Anabel, con quien hemos debatido todo lo debatible y se ha convertido en una querida amiga. También a Rafa y al León, a María y a Silvia, a Rufo y a Hèctor. Gente maravillosa, interesante, inteligente, honesta, que me han enriquecido enormemente con su amistad.

Como director de Tesis, Antonio ha tenido la excepcional capacidad de ser un acompañante sabio, un consejero, más que un guía a quien seguir. Me ha dejado transitar los caminos que yo eligiera, apoyándome, acompañándome, nunca coartándome, pero bajándome a tierra de vez en cuando. Me ha permitido desarrollar mi estilo, sin imponerme el suyo, pero ofreciéndome alternativas cuando me he encontrado perdido. Siempre con una visión estratégica, en cada discusión me ha aportado la perspectiva global que recurrentemente he necesitado. Valoro enormemente esta actitud tan poco frecuente entre los directores de Tesis, y se lo agradezco de corazón. Ojalá en el futuro, si alguna vez me encuentro del *otro lado*, yo tenga la misma capacidad de respetar y valorar las capacidades de los demás, sin buscar réplicas forzadas de mí mismo.

Finalmente, Victoria Agazzi, Viqui, ha sido mi amor, mi cómplice y todo. No soy capaz de distinguir dónde acaba ella y comienzo yo. Crecimos juntos, enriqueciéndonos el uno al otro, apoyándonos y complementándonos. Esta Tesis es también suya, y suya es mi felicidad.

Todo ello, todos ustedes, me han hecho inmensamente rico. Un tipo de riqueza que, sin embargo, no se posee a costa de otros sino que con todos es compartida. Esta Tesis es la prueba. Lo confieso.

Facundo Muñoz. Burjassot, septiembre de 2012.

Índice general

1	Introducción	1
2	Geoestadística	7
2.1	Estacionariedad	9
2.2	Análisis del semivariograma	12
2.3	Representaciones espectrales	19
2.3.1	La función de correlación Exponencial en el espacio Euclídeo	32
2.4	Predicción geoestadística: Kriging	37
2.5	Conclusiones	40
3	Geoestadística basada en el coste	41
3.1	Proyecto piloto	42
3.1.1	Mapas acústicos mediante técnicas geoestadísticas	42
3.1.2	Presentación de los datos	49
3.1.3	Predicción geoestadística	50
3.1.4	Problema específico y motivación de la solución.	51
3.2	Distancia basada en el coste	54
3.2.1	Definición y propiedades	54
3.2.2	Relación con la distancia euclídea	57
3.2.3	Validez del modelo	58
3.3	Cálculo geográfico	60
3.3.1	Introducción	60

3.3.2	Localizaciones de predicción	61
3.3.3	Mapas de distancias	62
3.3.4	Incorporación de información adicional	65
3.3.5	Representación de resultados	65
3.4	Conclusiones	66
4	Implementación Computacional	69
4.1	Plataforma de trabajo	69
4.2	Datos iniciales	70
4.3	Localizaciones de predicción	71
4.4	Cálculo de las matrices de distancias	71
4.5	Información adicional	74
4.6	Cálculos geoestadísticos	76
4.6.1	Depuración de los datos	76
4.6.2	Modelo de regresión sobre la media del proceso	78
4.6.3	Modelo de semivariograma	78
4.6.4	Predicción Kriging	80
4.7	Esquema global del proceso	80
4.8	Aplicación al proyecto piloto	82
4.8.1	Comparación de ambos tipos de distancias	82
4.8.2	Modelo de regresión sobre la media	84
4.8.3	Análisis del semivariograma	86
4.8.4	Resultados de predicción	89
4.9	Conclusiones	91
5	Geoestadística Riemanniana	95
5.1	La Variedad Riemanniana basada en el coste	98
5.2	Inmersiones en otros espacios	102
5.2.1	Inmersión de Kuratowski	102
5.2.2	Inmersión de Fréchet	104
5.2.3	Inmersión de Banach	105
5.3	Definición Positiva	105

5.4	Funciones definidas positivas en espacios no Euclídeos	113
5.5	Conclusiones	118
6	Representación pseudo-Euclídea	121
6.1	Multidimensional Scaling	122
6.1.1	Disimilitudes y distancias	124
6.1.2	Classical Scaling	125
6.2	Representación pseudo-Euclídea	129
6.2.1	Inmersibilidad en un espacio pseudo-Euclídeo	131
6.2.2	Fórmulas adicionales	137
6.2.3	Ejemplos	142
6.2.4	Relación con la Relatividad Especial	144
6.3	Definición Positiva en los espacios pseudo-Euclídeos	145
6.3.1	Resultados parciales sobre la función $M(\rho)$	152
6.4	Cálculo de funciones de densidad espectral	159
6.4.1	Densidad espectral de la función de correlación constante en $E_{(k,d-k)}$	160
6.4.2	Densidad espectral de una función de correlación isotrópica en $E_{(2,1)}$	161
6.4.3	Densidad espectral de la función de correlación exponencial en $E_{(2,1)}$	168
6.4.4	Expresiones alternativas de la función de correlación expo- nencial en $E_{(2,1)}$	170
6.5	Conclusiones	174
7	Reparametrización de la matriz de covarianzas	179
7.1	Descomposición de Cholesky	182
7.2	Conclusiones	185
8	Aproximaciones markovianas de campos Matérn	187
8.1	Vínculo entre GFs y GMRFs	188
8.2	Estructura de correlaciones en la región de Curriero	191
8.3	Conclusiones	195

9 Conclusiones y líneas futuras de trabajo	199
9.1 Conclusiones generales	199
9.2 Conclusiones específicas	200
9.3 Líneas abiertas de investigación	205
Apéndices	
A Deducción de las ecuaciones del Kriging	211
A.1 Kriging Simple	213
A.2 Kriging Ordinario	214
A.3 Inversión del sistema	216
B Glosario	219
C Implementación del cálculo de distancias basadas en el coste	225
D Adaptación del paquete <code>geoR</code>	237
E Geometría Riemanniana	255
E.1 Variedades Diferenciables	255
E.2 Espacio Tangente	258
E.3 Variedades Riemannianas	260
F Sistema de coordenadas pseudo-hiperesféricas del espacio pseudo-Euclídeo $E_{(k,d-k)}$	263
F.1 Puntos de cuadrado positivo $(\mathbf{x}, \mathbf{x}) > 0$	266
F.2 Caso $k = 1$	281
F.3 Puntos de cuadrado negativo $(\mathbf{x}, \mathbf{x}) < 0$	283
F.4 Rotaciones del espacio pseudo-Euclídeo	286
Bibliografía	291

Capítulo 1

Introducción

El germen de la presente Tesis consistió en un problema aplicado, de ingeniería, al que pensamos que la Estadística como disciplina puede contribuir de manera significativa.

Concretamente, se trata de la elaboración de mapas acústicos en entornos urbanos. Resuelto habitualmente de una manera determinista y aproximada, la valoración de la incertidumbre de los resultados es extremadamente deficiente en la mayoría de los casos reales.

Este problema, siendo de naturaleza espacial, se puede ver como un problema de predicción geoestadística, a partir de un conjunto de observaciones de campo. La dificultad radica en que el fenómeno se sitúa en un entorno urbano, que posee una importante heterogeneidad producida por los edificios, árboles y demás mobiliario urbano. La falta de homogeneidad en la región de interés hace inaplicables las técnicas geoestadísticas usuales.

Esto nos condujo a buscar metodologías geoestadísticas alternativas, más generales, que contemplaran la posibilidad de que la región de interés tuviera una naturaleza heterogénea.

Una adaptación interesante de la metodología geoestadística consiste en sustituir la distancia Euclídea entre localizaciones de la región por una medida de distancia que tenga en cuenta el grado de irregularidad local de la región.

Esta idea no es nueva. Diversos autores han implementado diferentes versiones de la misma. Muy frecuentemente en el contexto acuático, con islas o estuarios, etcétera. Es decir, cuando existen trozos de tierra que funcionan como obstáculos o barreras para la difusión del fenómeno. Así, han definido la llamada *distancia acuática* como la distancia mínima en la que un pez podría ir de un punto a otro a través del agua.

La implementación que nosotros proponemos en esta Tesis, en cambio, es original. Se basa en la construcción de una superficie *de coste*, que mide la heterogeneidad local de la región. A partir de ella se pueden calcular este tipo de distancias como la longitud de las rutas de mínimo coste entre localizaciones. Por este motivo la hemos denominado *distancia basada en el coste*.

Esta medida es más general que la distancia acuática. Mientras que las barreras absolutas (e.g., islas en el mar) se pueden representar a través de un coste infinito, hay infinitas situaciones intermedias que se pueden modelar con la superficie de coste. De este modo, en una región homogénea la distancia basada en el coste se reduce a la medida Euclídea, mientras que en general esta medida de distancia aumentaría continuamente en función del nivel de heterogeneidad.

El enfoque resultó tan interesante que utilizamos el problema de la elaboración de mapas acústicos en entornos urbanos como proyecto piloto. Implementamos los algoritmos para el cálculo de la distancia basada en el coste y adaptamos algunos algoritmos de geoestadística clásica —como la visualización y el ajuste de variogramas y la predicción Kriging— para utilizar distancias basadas en el coste. Los resultados fueron alentadores.

Sin embargo, un problema fundamental quedaba al margen de la implementación práctica de los algoritmos. Cuando se sustituye la distancia Euclídea por otro tipo de distancias se desmonta el sustento teórico que garantiza que la matriz de covarianzas de las observaciones sea definida positiva.

En el espacio Euclídeo, las funciones de correlación (o equivalentemente, los variogramas) *válidas* son completamente conocidas, en términos de representaciones espectrales. A lo largo de los años, diversos autores han encontrado múltiples familias paramétricas de funciones con distintas propiedades. Pero estas funciones pierden toda validez fuera del espacio Euclídeo.

Esto nos llevó a estudiar los aspectos teóricos de la definición positiva, con el objetivo último de encontrar (al menos) una familia de funciones válidas con la distancia basada en el coste. Sólo podemos presentar resultados parciales en este sentido y múltiples líneas futuras de investigación.

Desde el punto de vista matemático, el problema se puede modelar adecuadamente como una variedad con una métrica Riemanniana que depende directamente de la superficie de coste. La distancia basada en el coste resulta de forma natural como la distancia geodésica entre localizaciones de la variedad.

Por otra parte, esta variedad se puede *sumergir* en otros espacios más estructurados, con el objetivo de utilizar resultados conocidos sobre funciones definidas positivas. Naturalmente, la inmersión más interesante sería en un espacio Euclídeo. Pero demostramos que esto no es posible, en general, en ninguna cantidad, finita o infinita, de dimensiones.

Lo que sí es posible es obtener la mejor aproximación a una inmersión Euclídea en un número dado de dimensiones. Esto se conoce habitualmente como Multidimensional Scaling (MDS). Sin embargo, la aproximación produce un emborronamiento de las irregularidades de la región, que constituyen precisamente la motivación original del enfoque basado en el coste.

Aunque no es una solución satisfactoria, el MDS nos sirvió de inspiración para el desarrollo de otro tipo de inmersión, que constituye el segundo aporte original de esta Tesis. Se trata de la representación pseudo-Euclídea. Es una generalización del MDS que permite una representación *exacta* de una estructura de distancias no necesariamente Euclídea, en un espacio vectorial dotado de un *producto interno indefinido*.

Esta fue la línea de trabajo que nos pareció más promisoría en su momento, y en la que invertimos más tiempo y esfuerzos. Analizamos con todo detalle el

problema de la definición positiva en el espacio pseudo-Euclídeo, y estudiamos la representación espectral de algunas funciones candidatas.

Otra línea de trabajo explorada pero infructuosa consistió en la modificación del modelo, para trabajar con una reparametrización de la matriz de covarianzas que garantizara la definición positiva. Sin embargo, luego de una revisión de la literatura disponible, llegamos a la conclusión de que el enfoque en sí mismo es inadecuado para este caso, porque carece de un modelo subyacente que garantice la definición positiva para cualquier conjunto de puntos potencialmente elegible.

Por último, investigamos la conexión entre el enfoque basado en el coste y las aproximaciones markovianas de campos Matérn introducidas por [Lindgren et al. \(2011\)](#). Este último enfoque permite ajustar un campo aleatorio evitando la especificación de una función de correlación (y por ende, el problema de la definición positiva). Además es posible trabajar en superficies no planas, o incluso discontinuas.

La Tesis se estructura en dos grandes partes. En primer lugar abordamos el problema aplicado, comenzando con un breve repaso de la Teoría Geoestadística clásica en el [Capítulo 2](#). Ponemos algo de énfasis en la teoría de representación espectral, que necesitaremos especialmente en los [Capítulos 5 y 6](#).

En el [Capítulo 3](#) presentamos el proyecto piloto, la distancia basada en el coste y la metodología para su cálculo. Mientras que en el [Capítulo 4](#) describimos con detalle la implementación de la metodología basada en el coste, incluyendo las adaptaciones de los algoritmos geoestadísticos clásicos.

La segunda parte de la Tesis comienza con el [Capítulo 5](#), en el que situamos el problema de la Geoestadística basada en el coste en el contexto matemático de las variedades Riemannianas. Además discutimos la definición positiva en espacios no Euclídeos y evaluamos posibles inmersiones en espacios más estructurados.

En el [Capítulo 6](#) presentamos en detalle el trabajo desarrollado en torno a la representación pseudo-Euclídea.

Por último, en los [Capítulos 7 y 8](#) presentamos los enfoques alternativos correspondientes a la reparametrización de la matriz de covarianzas y la aproximación markoviana de campos Matérn.

Finalmente concluimos con una compilación de resultados y líneas futuras de investigación en el Capítulo 9.

En cuanto al estilo de redacción en lengua española, se han seguido los criterios de [Pérez Ortiz \(1999\)](#).

Capítulo 2

Geoestadística

La *Geoestadística* es el nombre genérico de una familia de técnicas estadísticas relativas al análisis de datos espaciales. Sus orígenes se suelen situar a principios de la década de 1950, en el trabajo del Ingeniero de Minas sudafricano Daniel G. Krige (Krige, 1951). Fue desarrollada y sistematizada posteriormente por Georges Matheron (Matheron, 1963). Algunas referencias clásicas en la materia son Journel y Huijbregts (1978); Cressie (1993) o Chilès y Delfiner (1999).

Los métodos geoestadísticos tienen una gran variedad de aplicaciones. La característica común que subyace en todas ellas es que las observaciones pueden ser vistas como una realización (usualmente parcial) de un proceso estocástico (o *campo aleatorio*, para destacar su naturaleza espacial)

$$\{Z(\mathbf{s}) : \mathbf{s} \in D\}, \tag{2.1}$$

donde D es un subconjunto fijo de \mathbb{R}^d tal que \mathbf{s} varía continuamente en la región D . Típicamente $d = 2$, y el índice \mathbf{s} representa una localización espacial en un *sistema de coordenadas* proyectado.

Los métodos geoestadísticos permiten modelar el proceso (2.1) con el objetivo de, a partir de un conjunto de observaciones $\{Z(\mathbf{s}_1), \dots, Z(\mathbf{s}_n)\}$, hacer

inferencia sobre los parámetros del modelo, efectuar contrastes de hipótesis y fundamentalmente realizar predicciones en localizaciones no observadas (ver Fig. 2.1).

La llamada *Primera Ley de la Geografía de Tobler* (Tobler, 1970) postula que «*Todo está relacionado con todo, pero las cosas cercanas están más relacionadas que las distantes*». En geoestadística, esta propiedad se modela a través de la *función de covarianza* o *covariograma*, que determina, para cada par de localizaciones $\mathbf{s}_1, \mathbf{s}_2 \in D$, la covarianza entre las correspondientes variables aleatorias:

$$C(\mathbf{s}_1, \mathbf{s}_2) = \mathbb{C}[Z(\mathbf{s}_1), Z(\mathbf{s}_2)].$$

Para todo n finito, todo conjunto de puntos $\mathbf{s}_1, \dots, \mathbf{s}_n \in D$ y todo conjunto de coeficientes reales a_1, \dots, a_n la combinación lineal $\sum_{i=1}^n a_i Z(\mathbf{s}_i)$ define una variable aleatoria, y por tanto debe ser

$$\mathbb{V}\left[\sum_{i=1}^n a_i Z(\mathbf{s}_i)\right] = \sum_{i,j=1}^n a_i a_j C(\mathbf{s}_i, \mathbf{s}_j) \geq 0. \quad (2.2)$$

Dicho de otra manera, la función de covarianza C de un proceso espacial bien definido debe verificar necesariamente la condición (2.2), en cuyo caso se denomina *definida positiva* en $D \times D$.

La condición de definida positiva, además de necesaria es también suficiente. En efecto, dada una función de media $m(\mathbf{s}) = \mathbb{E}[Z(\mathbf{s})]$ y una función definida positiva $C(\cdot, \cdot)$, existe un campo aleatorio Gaussiano con esta función de covarianza y aquella función de media. Basta tomar la distribución conjunta de $(Z(\mathbf{s}_1), \dots, Z(\mathbf{s}_n))$ como normal multivariante con media $(m(\mathbf{s}_1), \dots, m(\mathbf{s}_n))$ y matriz de covarianza con elemento i, j $K(\mathbf{s}_i, \mathbf{s}_j)$. Esta familia de distribuciones finito-dimensionales para todos los posibles conjuntos de puntos $\mathbf{s}_1, \dots, \mathbf{s}_n \in D$ satisfacen las condiciones de consistencia de Kolmogorov, y por tanto definen un campo aleatorio con estas distribuciones marginales (Gihman y Skorohod, 1974).

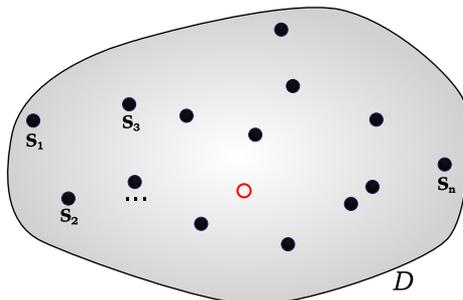


Figura 2.1. Predicción en una localización no observada dentro de la región de trabajo D .

2.1 Estacionariedad

Para hacer predicciones a partir de una única realización parcial de un proceso estocástico, es necesario asumir alguna hipótesis de regularidad. El proceso se suele descomponer en una parte no estocástica que recoge la variación a gran escala, y otra componente que recoge el comportamiento local, o variación a pequeña escala

$$Y(\mathbf{s}) = \mu(\mathbf{s}) + Z(\mathbf{s}).$$

De este modo, el comportamiento global está determinado (aunque sea desconocido) y se puede estimar mediante un modelo de regresión adecuado. Por su parte, sobre la variabilidad local se puede asumir la hipótesis de *estacionariedad débil*, o *de segundo orden*, que establece que la media del proceso es constante en toda la región y que la función de covarianza es invariante ante traslaciones; es decir, que depende sólo de la configuración relativa de las localizaciones. A veces es posible asumir la hipótesis adicional de *Isotropía*, que establece que la función de covarianza es también invariante ante rotaciones; es decir, que depende únicamente de la distancia que separa las localizaciones, independientemente de la dirección.

A lo largo de todo este trabajo se asumirán las hipótesis de estacionariedad débil e isotropía, que se formalizan a continuación.

Definición 2.1 (Procesos débilmente estacionarios e isotrópicos; covariograma).
Un proceso estocástico $\{Z(\mathbf{s}) : \mathbf{s} \in D\}$ que satisface:

$$\begin{aligned} i) \quad & \mathbb{E}[Z(\mathbf{s})] = \mu, \quad \forall \mathbf{s} \in D \\ ii) \quad & \mathbb{C}[Z(\mathbf{s}_1), Z(\mathbf{s}_2)] = C(\mathbf{s}_1 - \mathbf{s}_2), \quad \forall \mathbf{s}_1, \mathbf{s}_2 \in D \end{aligned}$$

se dice *débilmente estacionario*, o que satisface la hipótesis de *estacionariedad de segundo orden*. La función $C(\cdot)$ se llama *covariograma*.

Si además $C(\mathbf{s}_1 - \mathbf{s}_2)$, depende sólo de $\|\mathbf{s}_1 - \mathbf{s}_2\|$, entonces el covariograma $C(\cdot)$ se dice *isotrópico*.

Nota. Es habitual extender el adjetivo de *isotrópico* al proceso. En cualquier caso, sólo tiene sentido hablar de isotropía en el marco de procesos estacionarios.

Propiedades

A continuación se demuestran algunas consecuencias inmediatas de la hipótesis de estacionariedad débil.

1. Homocedasticidad

Como la función de covarianza es invariante ante traslaciones, en particular la varianza es constante y coincide con el valor de la función de covarianza en el origen:

$$\mathbb{V}[Z(\mathbf{s})] = \mathbb{C}[Z(\mathbf{s}), Z(\mathbf{s})] \stackrel{ii)}{=} C(\mathbf{0}) =: \sigma^2, \quad \forall \mathbf{s} \in D.$$

2. Estacionariedad de los incrementos

La varianza de los incrementos del proceso es también invariante ante traslaciones. Esto es, que se puede expresar en términos de una función γ que depende sólo de la configuración relativa de las localizaciones:

$$\begin{aligned} \mathbb{V}[Z(\mathbf{s}_1) - Z(\mathbf{s}_2)] &= \mathbb{V}[Z(\mathbf{s}_1)] + \mathbb{V}[Z(\mathbf{s}_2)] - 2\mathbb{C}[Z(\mathbf{s}_1), Z(\mathbf{s}_2)] \\ &= 2(\sigma^2 - C(\mathbf{s}_1 - \mathbf{s}_2)) =: 2\gamma(\mathbf{s}_1 - \mathbf{s}_2), \quad \forall \mathbf{s}_1, \mathbf{s}_2 \in D. \end{aligned}$$

La función $\gamma(\mathbf{h}) = \sigma^2 - C(\mathbf{h})$ se denomina *semivariograma* y proporciona una perspectiva diferente de las propiedades de segundo orden del proceso subyacente.

Nota. En virtud de *i*), también se puede escribir:

$$\mathbb{V}[Z(\mathbf{s}_1) - Z(\mathbf{s}_2)] = \mathbb{E} \left[(Z(\mathbf{s}_1) - Z(\mathbf{s}_2))^2 \right] - \underbrace{\mathbb{E} [Z(\mathbf{s}_1) - Z(\mathbf{s}_2)]^2}_{\stackrel{i)}{=} (\mu - \mu)^2 = 0}$$

$$\mathbb{E} \left[(Z(\mathbf{s}_1) - Z(\mathbf{s}_2))^2 \right] = 2\gamma(\mathbf{s}_1 - \mathbf{s}_2), \quad \forall \mathbf{s}_1, \mathbf{s}_2 \in D.$$

Por lo que el semivariograma también representa las medias cuadráticas de los incrementos del proceso.

3. Procesos estacionarios e isotrópicos

En caso de que el covariograma sea isotrópico, entonces él mismo y el semivariograma son funciones de una variable real no negativa h , que representa la distancia entre las localizaciones correspondientes:

$$\left. \begin{array}{l} \mathbb{V}[Z(\mathbf{s}_1) - Z(\mathbf{s}_2)] = 2\gamma(h) \\ \mathbb{C}[Z(\mathbf{s}_1), Z(\mathbf{s}_2)] = C(h) \end{array} \right\} \quad h = \|\mathbf{s}_1 - \mathbf{s}_2\|, \quad \forall \mathbf{s}_1, \mathbf{s}_2 \in D.$$

La Propiedad 2 conduce a la siguiente definición.

Definición 2.2 (Semivariograma). La semivarianza de los incrementos de un proceso estocástico $\{Z(\mathbf{s}) : \mathbf{s} \in D\}$ débilmente estacionario es una función invariante ante traslaciones que se denomina *semivariograma*:

$$\gamma(\mathbf{h}) := \frac{1}{2} \mathbb{V}[Z(\mathbf{s} + \mathbf{h}) - Z(\mathbf{s})], \quad \forall \mathbf{s} \in D.$$

Además, se dice que el semivariograma es *isotrópico* si es también invariante ante rotaciones. En ese caso, la función depende únicamente del módulo del vector \mathbf{h} .

Aunque tanto el semivariograma como el covariograma proporcionan sólo perspectivas diferentes de las propiedades de segundo orden del proceso, el semivariograma es preferido en los análisis geoestadísticos aplicados.

En primer lugar, el semivariograma está bien definido aunque no haya homocedasticidad. Por tanto, se puede trabajar con la hipótesis menos restrictiva de estacionariedad de los incrementos, en lugar de *ii*). Entonces se habla de un proceso *intrínsecamente estacionario*. Por otra parte, en este contexto, la función de covarianza depende de la media del proceso (pues $\mathbb{C}[Z(\mathbf{s}_1), Z(\mathbf{s}_2)] = \mathbb{E}[(Z(\mathbf{s}_1) - \mu)(Z(\mathbf{s}_2) - \mu)]$); de modo que la incertidumbre en la estimación de μ se propaga a la estimación de la función de covarianza. En cambio, la variabilidad de los incrementos es independiente de la media del proceso, lo que es más conveniente a los efectos de la estimación. Véase, por ejemplo, Gneiting et al. (2001); Schlather y Gneiting (2006) para más detalles sobre la relación entre los semivariogramas y las funciones de covarianza.

En los capítulos teóricos de esta Tesis (5, 6, 7 y 8), sin embargo, nos referiremos preferentemente al covariograma o a la función de correlación.

2.2 Análisis del semivariograma

En el enfoque clásico de la Geoestadística es necesario estimar el semivariograma del proceso a partir de los datos observados. Una forma sencilla de hacerlo es por el método de los momentos, que se basa en las medias cuadráticas muestrales de los incrementos del proceso (ver nota de la Propiedad 2). Fijado $h > 0$,

$$2\hat{\gamma}(h) = \frac{1}{\|N(h)\|} \sum_{N(h)} (Z(\mathbf{s}_1) - Z(\mathbf{s}_2))^2,$$

donde $N(h) = \{\{\mathbf{s}_i, \mathbf{s}_j\} : \|\mathbf{s}_i - \mathbf{s}_j\| = h, i, j = 1, \dots, n\}$ es el conjunto de pares de localizaciones separadas entre sí por una distancia h , y $\|N(h)\|$ es su cardinal.

En la práctica es imposible tener varios pares de observaciones a todas las posibles distancias h , de modo que se suele hacer una partición del espectro

de distancias en unos cuantos intervalos y se obtiene un valor estimado del semivariograma para cada uno. El gráfico resultante se denomina *semivariograma empírico*. La elección del número de intervalos de la partición es un problema análogo al que se presenta en la construcción de un histograma y se ha de atender al mismo criterio: ni demasiado pocos, para poder apreciar las diferencias, ni demasiados intervalos con pocas observaciones en los que la variabilidad dominante sea la del error del estimador (ver Fig. 2.2). [Journel y Huijbregts \(1978\)](#) sugieren, como regla práctica, asegurarse de que cada intervalo contenga al menos 30 pares de observaciones.

Una objeción a este estimador es su inestabilidad ante la presencia de valores extremos, por lo que se han propuesto diferentes estimadores robustos que o bien introducen un factor de corrección del sesgo, o hacen uso de la mediana ([Cressie y Hawkins, 1980](#)). Sin embargo, [Curriero et al. \(2002\)](#) compara este estimador clásico con otro tipo de estimadores llamados *no-ergódicos* mostrando que no existe una clara ventaja en su uso, si se toma en cuenta la efectividad en la predicción.

El semivariograma empírico proporciona, entonces, unas cuantas estimaciones de valores de la función de semivariograma $\gamma(\cdot)$ para algunos valores de distancia. Con esta información hay que determinar la función completa, teniendo en cuenta ciertas restricciones teóricas. Ya hemos visto al inicio del capítulo que una condición necesaria y suficiente para que una función de covarianza sea válida es la definición positiva (2.2). Dada la estrecha relación entre la función de covarianza y el variograma, esta condición induce inmediatamente la siguiente, para la función de semivariograma.

Definición 2.3 (Condición de negatividad). Una función de semivariograma es *condicionalmente definida negativa* si

$$\sum_{i=1}^m \sum_{j=1}^m a_i a_j \gamma(\mathbf{s}_i - \mathbf{s}_j) \leq 0,$$

para cualquier número finito de localizaciones espaciales $\{\mathbf{s}_i : i = 1, \dots, m\}$ y números reales $\{a_i : i = 1, \dots, m\}$ verificando $\sum_{i=1}^m a_i = 0$.

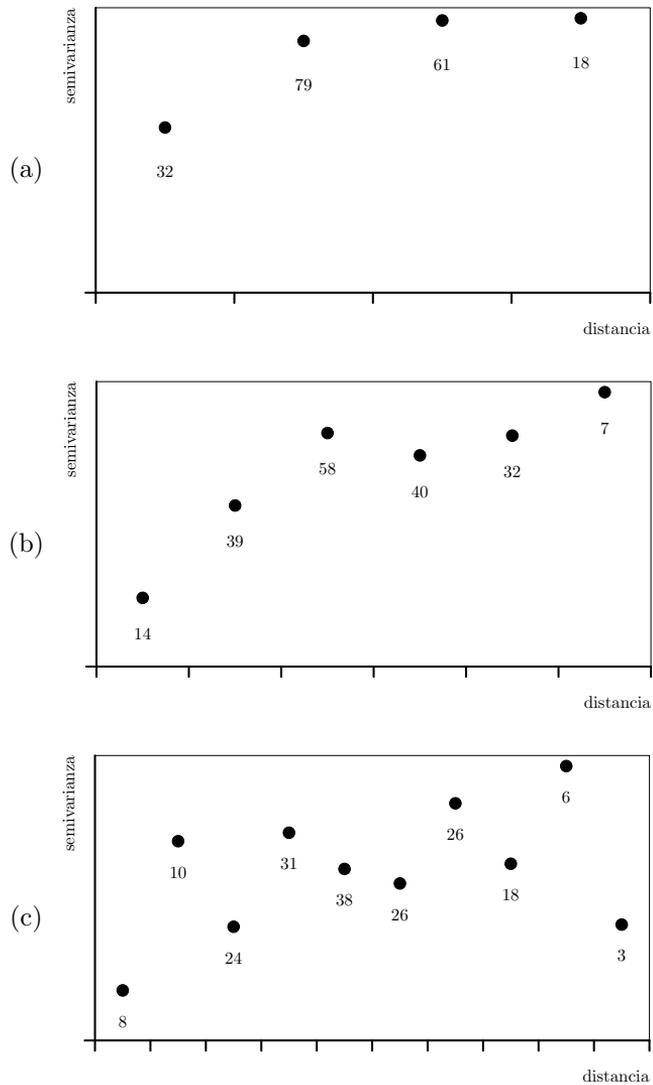


Figura 2.2. Semivariogramas empíricos. (a) Con pocos intervalos de partición se pierden los detalles; (c) con demasiados intervalos el error de estimación domina el gráfico. Hay que encontrar un equilibrio, como en (b).

Esta condición implica restricciones complejas en la estructura de correlaciones para asegurar que, para cualquier entero n , conjunto de localizaciones \mathbf{s}_i y constantes reales a_i sumando cero, la variable aleatoria definida por la combinación lineal $\sum_{i=1}^n a_i Z(\mathbf{s}_i)$ tenga una varianza no negativa.

En el espacio euclídeo \mathbb{R}^d el problema está completamente resuelto gracias al Teorema de Bochner (Bochner 1955; Teo. 2.4), que establece que las funciones de covariograma válidas son exactamente las funciones características de distribuciones d -dimensionales simétricas.

En la práctica (Le y Zidek, 2006) existe un conjunto estándar de familias de funciones paramétricas que aseguran la condición de condicionalmente definido negativo para el semivariograma $\gamma(\cdot)$. Los parámetros tienen una interpretación en términos de las dos características principales de un semivariograma:

- i)* su comportamiento en el origen;
- ii)* la existencia de una cota superior a su incremento.

El comportamiento en el origen está relacionado con la continuidad y regularidad del proceso. Notar que por la definición de semivariograma,

$$\gamma(0) = \frac{1}{2} \mathbb{V} [Z(\mathbf{s}) - Z(\mathbf{s})] = 0,$$

sin embargo, en general:

$$\lim_{h \rightarrow 0} \gamma(h) = \tau^2 \geq 0.$$

Al valor τ^2 se le conoce como *efecto pepita* y puede ser consecuencia de una discontinuidad en el proceso, del error en las mediciones, o de ambos simultáneamente. En determinadas aplicaciones se puede asumir la hipótesis de ausencia de efecto pepita. Sin embargo, suele existir una variabilidad intrínseca en las mediciones de fenómenos ambientales que se modela a través de este parámetro.

La segunda característica se refiere al comportamiento en distancias grandes. En general, cuando h crece, la correlación entre dos variables $Z(\mathbf{s} + \mathbf{h})$ y $Z(\mathbf{s})$

disminuye y eventualmente desaparece (recordar la Ley de Tobler mencionada al inicio de la Sección 2.1). Así,

$$\lim_{h \rightarrow \infty} \gamma(h) = \sigma^2 - \lim_{h \rightarrow \infty} C(h) = \sigma^2,$$

llamándose a σ^2 el *alféizar* del semivariograma.

El semivariograma puede alcanzar el alféizar para un cierto valor $\phi < \infty$, o bien acercarse asintóticamente. En el primer caso, al valor ϕ se le llama *rango* y representa la distancia a partir de la cual ya no existe correlación entre las variables. En el segundo caso, se define el *rango efectivo* como la distancia ϕ a partir de la cual el semivariograma alcanza el 95 % del alféizar (ver Fig. 2.3).

En la Figura 2.4 se muestra esquemáticamente la forma que adoptan algunas de las familias estándar más utilizadas, cuyas expresiones se especifican a continuación, en términos de los tres parámetros característicos: alféizar σ^2 ; pepita τ^2 y rango (efectivo) ϕ . Se omite especificar el valor en el origen, $\gamma(0) = 0$, que es común en todos los casos.

Efecto

pepita puro

$$\gamma(h) \equiv \tau^2 \geq 0, \quad h > 0.$$

Modelo

Exponencial

$$\gamma(h) = \tau^2 + (\sigma^2 - \tau^2) \left(1 - e^{-3\frac{h}{\phi}}\right), \quad h > 0.$$

Modelo

Gaussiano

$$\gamma(h) = \tau^2 + (\sigma^2 - \tau^2) \left(1 - e^{-3\left(\frac{h}{\phi}\right)^2}\right), \quad h > 0.$$

Modelo

Esférico

$$\gamma(h) = \begin{cases} \tau^2 + (\sigma^2 - \tau^2) \left(\frac{3h}{2\phi} - \frac{h^3}{2\phi^3}\right), & 0 < h \leq \phi, \\ \sigma^2, & h \geq \phi. \end{cases}$$

Modelo

Matérn

$$\gamma_\nu(h) = \tau^2 + (\sigma^2 - \tau^2) \left(1 - \frac{(2\sqrt{\nu}h/\phi)^\nu}{2^{\nu-1}\Gamma(\nu)} K_\nu\left(\frac{2\sqrt{\nu}h}{\phi}\right)\right), \quad h > 0;$$

donde K_ν es la función de Bessel modificada de segunda especie de orden ν (Abramowitz y Stegun, 1972).

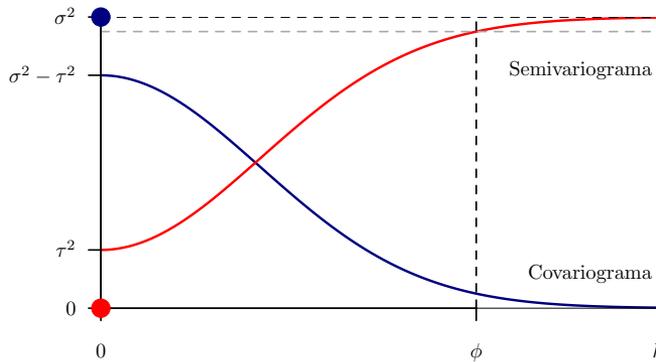


Figura 2.3. Modelo de semivariograma típico, sus parámetros característicos y el covariograma correspondiente.

La familia Matérn (Matérn, 1986) es una clase más general de funciones que depende de un parámetro adicional ν que regula la suavidad del proceso estocástico. Contiene como caso particular a la familia Exponencial ($\nu = \frac{1}{2}$) y a los modelos Gaussianos ($\nu \rightarrow \infty$). Cuando $\nu = 1$ se obtiene la familia Whittle, que no ha sido expuesta aquí. Más información sobre la familia Matérn puede encontrarse en Handcock y Stein (1993), Stein (1999) o Guttorp y Gneiting (2006), y en general sobre los modelos de semivariograma y sus propiedades en Diggle y Ribeiro (2007), en Schlather (1999) o en Journel y Huijbregts (1978).

La estimación de la estructura de correlaciones se realiza entonces en términos de una función de semivariograma cuyos parámetros se ajustan a partir de las observaciones. Existen varios métodos para estimar los parámetros de la función de semivariograma: máxima verosimilitud (Kitanidis, 1983); máxima verosimilitud restringida; mínima norma cuadrática (Rao, 1972, 1979); mínimos cuadrados y mínimos cuadrados generalizados (Cressie, 1993). Un enfoque Bayesiano del problema permite tener en cuenta la incertidumbre en los parámetros del vario-

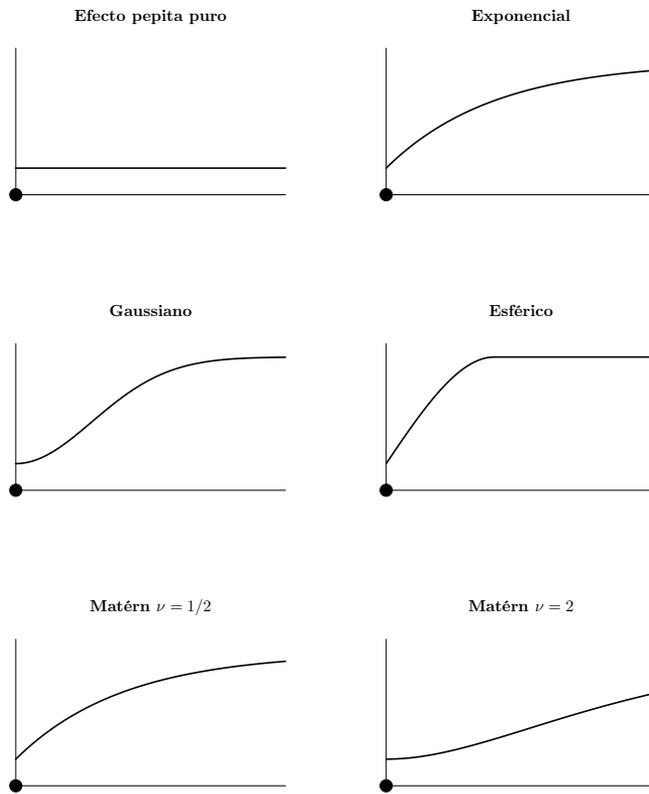


Figura 2.4. Esquemas de algunos modelos de semivariogramas.

grama, considerándolos parámetros desconocidos adicionales e integrándolos en un modelo jerárquico (Banerjee et al., 2004).

Nuestra elección para la parte aplicada de esta Tesis ha sido el método de mínimos cuadrados, ajustando simultáneamente la función de media $\mu(\mathbf{s})$ potencialmente dependiente de covariables adicionales y los parámetros de la función de semivariograma $\gamma(r)$. Esta elección está justificada por simplicidad. En esta Tesis nos ocupamos de la metodología de modelización de irregularidades, cuestión que se puede tratar desde un punto de vista tanto clásico como Bayesiano.

2.3 Representaciones espectrales

Los métodos espectrales constituyen una herramienta muy poderosa para el estudio de los campos aleatorios. En el Análisis de Fourier resulta más natural considerar funciones complejas, aunque en la práctica geoestadística estemos habitualmente interesados en campos reales.

Un campo aleatorio Z en \mathbb{R}^d se dice complejo si $Z(\mathbf{x}) = U(\mathbf{x}) + iV(\mathbf{x})$, donde U y V son campos aleatorios reales. Si U y V son *conjuntamente* débilmente estacionarios, esto es que U y V lo son y que $\mathbb{C}[U(\mathbf{x}_i), V(\mathbf{x}_j)]$ dependa sólo de $\mathbf{x}_i - \mathbf{x}_j$, entonces Z es débilmente estacionario, y se define

$$\begin{aligned} C(\mathbf{h}) &= \mathbb{C} \left[Z(\mathbf{x} + \mathbf{h}), \overline{Z(\mathbf{x})} \right] \\ &= \mathbb{C} [U(\mathbf{x} + \mathbf{h}), U(\mathbf{x})] + \mathbb{C} [V(\mathbf{x} + \mathbf{h}), V(\mathbf{x})] \\ &\quad + i \left(\mathbb{C} [U(\mathbf{x}), V(\mathbf{x} + \mathbf{h})] - \mathbb{C} [U(\mathbf{x} + \mathbf{h}), V(\mathbf{x})] \right) \end{aligned}$$

como la función de covarianza de Z . De esta manera, se cumple $C(-\mathbf{h}) = \overline{C(\mathbf{h})}$, y para todo conjunto de complejos $\alpha_1, \dots, \alpha_n$,

$$\sum_{i,j=1}^n \alpha_i \bar{\alpha}_j C(\mathbf{x}_i - \mathbf{x}_j) \geq 0,$$

ya que el lado izquierdo es igual a $\mathbb{E} \left[\left| \sum_{i=1}^n \alpha_i Z(\mathbf{x}_i) - \mathbb{E} \left[\sum_{i=1}^n \alpha_i Z(\mathbf{x}_i) \right] \right|^2 \right]$. Una función C que satisface esta condición para todo n , todo conjunto de localizaciones $\mathbf{x}_1, \dots, \mathbf{x}_n \in D$ y todo conjunto de complejos $\alpha_1, \dots, \alpha_n$ se denomina función definida positiva compleja en D .

Como ejemplo de campo aleatorio complejo, tómense $\boldsymbol{\omega}_1, \dots, \boldsymbol{\omega}_n \in \mathbb{R}^d$ y sean Z_1, \dots, Z_n variables aleatorias complejas con media 0, $\mathbb{E} [Z_i \bar{Z}_j] = 0$ si $i \neq j$ y $\mathbb{E} [|Z_i|^2] = f_i$. Considérese entonces,

$$Z(\mathbf{x}) = \sum_{k=1}^n Z_k e^{i\boldsymbol{\omega}'_k \mathbf{x}}, \quad (2.3)$$

de modo que Z_k es la *amplitud* aleatoria compleja de Z en la frecuencia $\boldsymbol{\omega}_k$. Entonces Z es un campo débilmente estacionario complejo en \mathbb{R}^d con función de covarianza

$$C(h) = \sum_{k=1}^n f_k e^{i\boldsymbol{\omega}'_k h}.$$

La Ecuación (2.3) es un ejemplo de representación espectral de un campo aleatorio. Tomando límites en L^2 de forma apropiada, se pueden obtener representaciones espectrales de todos los campos aleatorios débilmente estacionarios y continuos en media cuadrática (Yaglom, 1987). Dicho de otra manera, todos los campos aleatorios de este tipo son, en este sentido, combinaciones lineales de exponenciales complejas con amplitudes aleatorias incorreladas.

Este proceso de límite exige considerar medidas complejas aleatorias. Esto es, aplicaciones M que asignan variables aleatorias complejas a los subconjuntos de Borel en \mathbb{R}^d . Ver, por ejemplo, Gihman y Skorohod (1974) para los detalles matemáticos.

De este modo, la representación espectral de un campo Z se puede escribir

$$Z(\mathbf{x}) = \int_{\mathbb{R}^d} e^{i\boldsymbol{\omega}' \mathbf{x}} M(d\boldsymbol{\omega}).$$

Si además, para todos los subconjuntos de Borel Δ y para cierta medida positiva finita F ,

$$\begin{aligned}\mathbb{E}[M(\Delta)] &= 0 \\ \mathbb{E}[|M(\Delta)|^2] &= F(\Delta)\end{aligned}$$

y para cualesquiera subconjuntos de Borel disjuntos Δ_1, Δ_2 ,

$$\mathbb{E}\left[M(\Delta_1)\overline{M(\Delta_2)}\right] = 0,$$

entonces la función de covarianza de Z tiene la siguiente representación espectral (Stein, 1999, §2.5):

$$C(\mathbf{x}) = \int_{\mathbb{R}^d} e^{i\boldsymbol{\omega}'\mathbf{x}} F(d\boldsymbol{\omega}).$$

La medida F se llama *medida espectral* o *espectro* de Z .

El siguiente teorema, central en la Teoría de representación espectral, caracteriza todas las funciones de covarianza válidas para algún campo aleatorio débilmente estacionario y continuo en media cuadrática como aquellas que admiten una representación espectral de este tipo.

Teorema 2.4. *Bochner (1959)*

Una función real $R(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^d$ es definida positiva si y sólo si se puede representar como

$$R(\mathbf{x}) = \int_{\mathbb{R}^d} e^{i\boldsymbol{\omega}'\mathbf{x}} F(d\boldsymbol{\omega}), \quad (2.4)$$

donde F es una medida positiva finita.

Demostración.

La siguiente demostración es una ampliación de la que se puede encontrar en Gihman y Skorohod (1974, p.208).

Es inmediato comprobar la suficiencia, ya que dados n puntos $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^d$ y n coeficientes complejos $\alpha_1, \dots, \alpha_n$,

$$\begin{aligned}
 \sum_{i,j=1}^n \bar{\alpha}_i \alpha_j R(\mathbf{x}_j - \mathbf{x}_i) &= \sum_{i,j=1}^n \bar{\alpha}_i \alpha_j \int_{\mathbb{R}^d} e^{i\boldsymbol{\omega}'(\mathbf{x}_j - \mathbf{x}_i)} F(d\boldsymbol{\omega}) \\
 &= \int_{\mathbb{R}^d} \sum_{i,j=1}^n \bar{\alpha}_i e^{-i\boldsymbol{\omega}'\mathbf{x}_i} \alpha_j e^{i\boldsymbol{\omega}'\mathbf{x}_j} F(d\boldsymbol{\omega}) \\
 &= \int_{\mathbb{R}^d} \overline{\left(\sum_{i=1}^n \alpha_i e^{i\boldsymbol{\omega}'\mathbf{x}_i} \right)} \left(\sum_{j=1}^n \alpha_j e^{i\boldsymbol{\omega}'\mathbf{x}_j} \right) F(d\boldsymbol{\omega}) \\
 &= \int_{\mathbb{R}^d} \left| \sum_i \alpha_i e^{i\boldsymbol{\omega}'\mathbf{x}_i} \right|^2 F(d\boldsymbol{\omega}) \geq 0.
 \end{aligned}$$

En este último paso se usa el hecho de que la medida F es positiva y finita. Esto muestra que toda función de la forma (2.4) es una función de correlación válida de cierto campo aleatorio débilmente estacionario y continuo en media cuadrática.

Para la necesidad, tenemos que ver que cualquier función continua y definida positiva admite la representación (2.4). Por la condición de definición positiva, para cualquier función arbitraria g , integrable en \mathbb{R}^d se verifica

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} R(\mathbf{x} - \mathbf{y}) \overline{g(\mathbf{x})} g(\mathbf{y}) d\mathbf{x} d\mathbf{y} \geq 0.$$

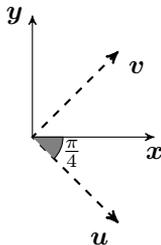
Consideremos $g(\mathbf{x}) = \exp(-\|\mathbf{x}\|^2/2N + i\mathbf{x}'\mathbf{z})$, para ciertos $N \geq 0$ y $\mathbf{z} \in \mathbb{R}^d$. Entonces,

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} R(\mathbf{x} - \mathbf{y}) e^{-\frac{\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2}{2N}} e^{-i(\mathbf{x} - \mathbf{y})'\mathbf{z}} d\mathbf{x} d\mathbf{y} \geq 0.$$

Aplicamos el siguiente cambio de variable ortogonal

$$\mathbf{u} = \frac{\mathbf{x} - \mathbf{y}}{\sqrt{2}}; \quad \mathbf{v} = \frac{\mathbf{x} + \mathbf{y}}{\sqrt{2}}.$$

Notar que se trata de una rotación horaria de ángulo $\pi/4$, es decir, una transformación ortogonal, que conserva distancias y ángulos (i.e., el producto interno):



Matricialmente se puede escribir $\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} = \frac{\sqrt{2}}{2} \begin{pmatrix} \mathbf{I}_d & -\mathbf{I}_d \\ \mathbf{I}_d & \mathbf{I}_d \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}$. La matriz del cambio es ortogonal y tiene jacobiano igual a 1.

Teniendo en cuenta que $\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$, resulta

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} R(\sqrt{2}\mathbf{u}) e^{-\frac{\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2}{2N}} e^{-i\sqrt{2}\mathbf{u}'z} d\mathbf{u}d\mathbf{v} \geq 0.$$

De este modo conseguimos separar variables,

$$\int_{\mathbb{R}^d} e^{-\frac{\|\mathbf{v}\|^2}{2N}} d\mathbf{v} \int_{\mathbb{R}^d} R(\sqrt{2}\mathbf{u}) e^{-\frac{\|\mathbf{u}\|^2}{2N}} e^{-i\sqrt{2}\mathbf{u}'z} d\mathbf{u} \geq 0.$$

El primer factor es positivo pues se reduce a la integral de Gauss:

$$\left(\int_{-\infty}^{\infty} e^{-\frac{x^2}{2N}} dx \right)^d = (2\pi N)^{d/2} > 0.$$

Por tanto, el segundo factor es no-negativo. Si lo reescribimos en términos de la variable $\boldsymbol{\omega} = \sqrt{2}\mathbf{u}$, de modo que $\|\mathbf{u}\|^2 = \|\boldsymbol{\omega}\|^2$, $d\mathbf{u} = \frac{1}{\sqrt{2}}d\boldsymbol{\omega}$, se tiene que la función

$$\tilde{R}_N(z) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} R(\boldsymbol{\omega}) e^{-\frac{\|\boldsymbol{\omega}\|^2}{4N}} e^{-i\boldsymbol{\omega}'z} d\boldsymbol{\omega}$$

es no-negativa. Además, \tilde{R}_N es la transformada de Fourier de la función integrable y continua $R(\mathbf{x})e^{-\frac{\|\mathbf{x}\|^2}{4N}}$ y también es diferenciable.

Tenemos que ver que \tilde{R}_N también es integrable. Por una parte, \tilde{R}_N es la transformada de Fourier de $R(\mathbf{x})e^{-\frac{\|\mathbf{x}\|^2}{4N}}$ y por su parte, $e^{-\varepsilon\frac{\|\boldsymbol{\omega}\|^2}{2}}$ es la transformada de Fourier de $\varepsilon^{-d/2}e^{-\frac{\|\mathbf{x}\|^2}{2\varepsilon}}$. Usando la igualdad de Parseval en la forma de producto interno $(\int f(x)\overline{g(x)}dx = \int \tilde{f}(\xi)\overline{\tilde{g}(\xi)}d\xi)$ se tiene:

$$\begin{aligned} \int_{\mathbb{R}^d} \tilde{R}_N(\boldsymbol{\omega})e^{-\varepsilon\frac{\|\boldsymbol{\omega}\|^2}{2}}d\boldsymbol{\omega} &= \int_{\mathbb{R}^d} \underbrace{R(\mathbf{x})}_{\leq R(0)} \underbrace{e^{-\frac{\|\mathbf{x}\|^2}{4N}}}_{\leq 1} \varepsilon^{-d/2}e^{-\frac{\|\mathbf{x}\|^2}{2\varepsilon}}d\mathbf{x} \\ &\leq R(0)\varepsilon^{-d/2} \underbrace{\int_{\mathbb{R}^d} e^{-\frac{\|\mathbf{x}\|^2}{2\varepsilon}}d\mathbf{x}}_{\left(\int_{-\infty}^{\infty} \varepsilon^{-1/2}e^{-\frac{x^2}{2\varepsilon}}\right)^d = (2\pi)^{d/2}}. \end{aligned}$$

De modo que para todo $\varepsilon > 0$ esta integral está acotada por una cantidad fija. Haciendo tender $\varepsilon \rightarrow 0$, por el Lema de Fatou,

$$\int_{\mathbb{R}^d} \tilde{R}_N(\boldsymbol{\omega})d\boldsymbol{\omega} \leq R(0)(2\pi)^{d/2},$$

y por tanto \tilde{R}_N es integrable.

Ahora estamos habilitados a utilizar la fórmula de inversión de la transformada de Fourier:

$$R(\mathbf{x})e^{-\frac{\|\mathbf{x}\|^2}{4N}} = \int_{\mathbb{R}^d} e^{i\mathbf{x}'\boldsymbol{\omega}}(2\pi)^{-d/2}\tilde{R}_N(\boldsymbol{\omega})d\boldsymbol{\omega} = \int_{\mathbb{R}^d} e^{i\mathbf{x}'\boldsymbol{\omega}}F_N(d\boldsymbol{\omega}),$$

donde

$$F_N(A) = \int_A (2\pi)^{-d/2}\tilde{R}_N(\boldsymbol{\omega})d\boldsymbol{\omega}.$$

Así, $\frac{R(\mathbf{x})}{R(0)}e^{-\frac{\|\mathbf{x}\|^2}{4N}}$ es la función característica de una cierta distribución F_N en \mathbb{R}^d y converge a una función continua cuando $N \rightarrow \infty$. Por tanto, en el límite, $R(\mathbf{x})/R(0)$ es también una función característica de una cierta medida F .

Además se tiene que F es única, por el teorema de la unicidad de una distribución con una función característica dada. \square

Si F tiene una densidad con respecto de la medida de Lebesgue, llamaremos a esta densidad la *densidad espectral* y la denotaremos generalmente por f . Cuando la densidad espectral existe, se puede calcular a partir de la función de correlación utilizando la fórmula usual de inversión de una transformada de Fourier d -dimensional:

$$f(\boldsymbol{\omega}) = (2\pi)^{-d} \int_{\mathbb{R}^d} R(\mathbf{x}) e^{-i\boldsymbol{\omega}'\mathbf{x}} d\mathbf{x}. \quad (2.5)$$

En este trabajo asumiremos que la densidad espectral existe y utilizaremos la Fórmula (2.5) para su cálculo.

En resumen, el Teorema de Bochner caracteriza la clase de todas las funciones de covarianza continuas en E_d como las transformadas de Fourier de todas las medidas positivas y finitas en \mathbb{R}^d .

En la práctica, sin embargo, es usual (y muy útil) trabajar con funciones de covarianza isotrópicas. Esto es, funciones que dependen de la diferencia vectorial entre los puntos únicamente a través de su magnitud, pero no de su dirección. Es decir, permitiéndonos un pequeño abuso de notación al denotar de la misma manera a la función de correlación vectorial y a la isotrópica,

$$\mathbb{C}[Z(\mathbf{x}_1), Z(\mathbf{x}_2)] = R(\mathbf{x}_2 - \mathbf{x}_1) = R(\|\mathbf{x}_2 - \mathbf{x}_1\|).$$

Por lo que nos interesa obtener una representación análoga para esta subclase de funciones.

Teorema 2.5. *Schoenberg (1938b)*

La clase $\mathcal{P}(E_d)$ de funciones reales definidas positivas en el espacio Euclídeo E_d es idéntica a la clase de funciones de la forma

$$R(h) = 2^{\frac{d-2}{2}} \Gamma(d/2) \int_0^\infty \frac{J_{\frac{d-2}{2}}(h\lambda)}{(h\lambda)^{\frac{d-2}{2}}} G(d\lambda), \quad (2.6)$$

donde J_ν es la función de Bessel de primera especie (ver 9.1.20 en [Abramowitz y Stegun, 1972](#)), y G es una medida finita en $[0, \infty)$.

La expresión a la derecha en (2.6) es conocida como la transformada de Hankel de orden $\frac{d-2}{2}$ de G .

Demostración.

La demostración a continuación es una elaboración propia a partir de la original de [Schoenberg \(1938b\)](#) y de formulaciones posteriores en [Gihman y Skorohod \(1974\)](#) y en [Stein \(1999\)](#).

Dada cualquier función de covarianza isotrópica R continua y definida positiva, por el Teorema de Bochner existe una medida positiva y finita F tal que para todo $\mathbf{x} \in \mathbb{R}^d$

$$R(\mathbf{x}) = \int_{\mathbb{R}^d} e^{i\mathbf{u}'\mathbf{x}} F(d\mathbf{u}).$$

Por la isotropía, $R(\mathbf{x})$ es constante sobre la superficie de la hipersfera B_h de radio h . Integrando la ecuación sobre dicha superficie se obtiene, en el lado izquierdo (Lema 2.6, a continuación),

$$R(h) \frac{2\pi^{d/2} h^{d-1}}{\Gamma(d/2)}.$$

Integrando el lado derecho y denotando por $ds(\mathbf{x})$ al elemento de área de la superficie esférica de radio h se obtiene

$$\begin{aligned} \int_{\partial B_h} \left(\int_{\mathbb{R}^d} e^{i\mathbf{u}'\mathbf{x}} F(d\mathbf{u}) \right) ds(\mathbf{x}) &= \int_{\mathbb{R}^d} \left(\int_{\partial B_h} e^{i\mathbf{u}'\mathbf{x}} ds(\mathbf{x}) \right) F(d\mathbf{u}) \\ &= \int_{\mathbb{R}^d} \left(\frac{1\pi h^{d/2}}{\|\mathbf{u}\|} \|\mathbf{u}\| J_{\frac{d-2}{2}}(h\|\mathbf{u}\|) \right) F(d\mathbf{u}). \end{aligned}$$

El cambio en el orden de integración está justificado por un argumento general en [Bochner \(1959, p. 309\)](#), mientras que el cálculo de la integral interior de efectúa en el Lema 2.7 (a continuación) mediante un cambio a coordenadas esféricas.

Igualando ambas expresiones y reordenando convenientemente se tiene

$$R(h) = 2^{\frac{d-2}{2}} \Gamma(d/2) \int_{\mathbb{R}^d} \frac{J_{\frac{d-2}{2}}(h\|\mathbf{u}\|)}{(h\|\mathbf{u}\|)^{\frac{d-2}{2}}} F(d\mathbf{u}).$$

Finalmente, como el integrando sólo depende de $\|\mathbf{u}\|$, hacemos el cambio de variable $\lambda = \|\mathbf{u}\|$ y definimos una medida G sobre el semieje $[0, \infty)$ mediante $G([a, b]) = F\{B_b \setminus B_a\}$, $0 \leq a \leq b$.

Con esto se tiene probado el teorema, puesto que G es claramente finita porque F lo es.

Sólo resta probar los dos lemas técnicos utilizados en el argumento. \square

Lema 2.6. *Integral de una función radial sobre la superficie de una d -esfera.*

Consideremos la esfera d -dimensional de radio h , $B_h = \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\| \leq h\}$. Se quiere calcular la integral de una función radial (i.e., que sólo depende de la distancia al origen) sobre su superficie $\partial B_h = \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\| = h\}$.

Denotando por $s(d\mathbf{x})$ al elemento de superficie, se tiene:

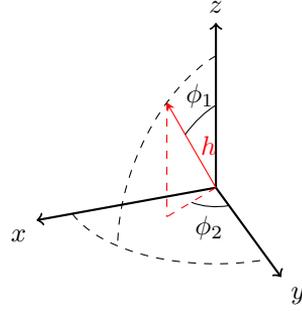
$$\begin{aligned} \int_{\partial B_h} f(h)s(d\mathbf{x}) &= f(h) \int_{\partial B_h} s(d\mathbf{x}) \\ &= f(h)h^{d-1} \int_{\partial B_1} s(d\mathbf{x}) \\ &= f(h)h^{d-1} \frac{2\pi^{d/2}}{\Gamma(d/2)}. \end{aligned} \tag{2.7}$$

Demostración. En el primer paso, $f(r)$ sale de la integral porque es constante sobre la superficie ∂B_h .

Para el segundo paso, utilizamos el Teorema de la Divergencia en coordenadas hiper-esféricas para ver que el área de una esfera de radio h en un espacio d -dimensional es h^{d-1} veces el área de la esfera unidad.

En efecto, las coordenadas hiper-esféricas generalizan las conocidas coordenadas esféricas representadas en el diagrama a espacios de dimensiones mayores de la siguiente manera,

$$\left\{ \begin{array}{l} x_1 = h \sin \phi_1 \sin \phi_2 \cdots \sin \phi_{d-2} \sin \phi_{d-1} \\ x_2 = h \sin \phi_1 \sin \phi_2 \cdots \sin \phi_{d-2} \cos \phi_{d-1} \\ x_3 = h \sin \phi_1 \sin \phi_2 \cdots \cos \phi_{d-2} \\ \vdots \\ x_{d-1} = h \sin \phi_1 \cos \phi_2 \\ x_d = h \cos \phi_1 \end{array} \right.$$



con $h \geq 0$ y $(\phi_1, \dots, \phi_{d-1}) \in \Phi = [0, \pi]^{d-2} \times [0, 2\pi)$.

El jacobiano de este cambio de variables es

$$J(h) = \begin{vmatrix} \sin \phi_1 \sin \phi_2 \cdots & h \cos \phi_1 \sin \phi_2 \cdots & \cdots \\ \sin \phi_1 \sin \phi_2 \cdots & h \cos \phi_1 \sin \phi_2 \cdots & \cdots \\ \vdots & \vdots & \vdots \\ \sin \phi_1 \cos \phi_2 & h \cos \phi_1 \cos \phi_2 & \cdots \\ \cos \phi_1 & -h \sin \phi_1 & \cdots \end{vmatrix} = h^{d-1} J(1),$$

porque h multiplica a $d - 1$ de las columnas del determinante.

El Teorema de la Divergencia establece que el flujo de un campo vectorial a través de una superficie cerrada S es igual a la integral volumétrica de su divergencia, en el interior de dicha superficie.

Consideremos el campo \mathbf{x} . Si $\hat{\mathbf{n}}$ es el vector unitario normal a la superficie, el flujo a través de ∂B_h es

$$\int_{\partial B_h} \mathbf{x} \cdot \hat{\mathbf{n}} s(d\mathbf{x}) = h \int_{\partial B_h} s(d\mathbf{x}),$$

puesto que $\mathbf{x} = h\hat{\mathbf{n}}$.

Por otra parte, la divergencia de \mathbf{x} en el punto \mathbf{x} es

$$\nabla \cdot \mathbf{x} = \frac{\partial x_1}{\partial x_1} + \cdots + \frac{\partial x_d}{\partial x_d} = d,$$

por tanto, cambiando a coordenadas esféricas,

$$\begin{aligned} \int_{B_h} \nabla \cdot \mathbf{x} \, d\mathbf{x} &= \int_{B_h} d \, d\mathbf{x} \\ &= \left(\int_0^h dr^{d-1} \, dr \right) \left(\int_{\Phi} J(1) \, d\phi_1 \cdots d\phi_{d-1} \right) \\ &= h^d \int_{\partial B_1} s(d\mathbf{x}). \end{aligned}$$

De la igualdad de estas dos cantidades sale el segundo paso en (2.7).

Finalmente, el último paso consiste en el cálculo de la superficie de la d -esfera unitaria, utilizando un ingenioso artilugio basado en las integrales Gamma y de Gauss.

Consideremos la cantidad

$$I = \int_{\mathbb{R}^d} e^{-\|\mathbf{x}\|^2} \, d\mathbf{x}.$$

En coordenadas hiperesféricas,

$$\begin{aligned} I &= \int_0^\infty \int_{\Phi} e^{-r^2} J(r) \, dr d\phi_1 \cdots d\phi_{d-1} \\ &= \left(\int_0^\infty r^{d-1} e^{-r^2} \, dr \right) \left(\int_{\Phi} J(1) \, d\phi_1 \cdots d\phi_{d-1} \right) \\ &= \frac{1}{2} \Gamma(d/2) \int_{\partial B_1} s(d\mathbf{x}). \end{aligned}$$

En el último paso se ha utilizado la Integral Gamma $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} \, dt$, con el cambio $t = r^2$ y $z = d/2$.

Por otra parte, en coordenadas cartesianas,

$$\begin{aligned} I &= \int_{\mathbb{R}^d} e^{-(x_1^2 + \cdots + x_d^2)} \, dx_1 \cdots dx_d \\ &= \prod_{i=1}^d \left(\int_{-\infty}^\infty e^{-x^2} \, dx \right) = \pi^{d/2}. \end{aligned}$$

En el último paso se ha utilizado que $G = \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$, pues mediante un cambio a coordenadas polares se tiene,

$$\begin{aligned} G^2 &= \left(\int_{-\infty}^{\infty} e^{-x^2} dx \right) \left(\int_{-\infty}^{\infty} e^{-y^2} dy \right) = \iint_{-\infty}^{\infty} e^{-(x^2+y^2)} dx dy \\ &= \int_0^{\infty} \int_0^{2\pi} e^{-r^2} r dr d\theta = 2\pi \left[\frac{-e^{-r^2}}{2} \right]_0^{\infty} = \pi. \end{aligned}$$

O bien, utilizando la integral de Gauss $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1$, con el cambio $x = y\sqrt{2}$.

Como se trata de la misma cantidad calculada de dos maneras diferentes, ambas expresiones son iguales, de donde se deduce la fórmula de la superficie de la d -esfera unidad:

$$\int_{\partial B_1} s(d\mathbf{x}) = \frac{2\pi^{d/2}}{\Gamma(d/2)}.$$

□

Lema 2.7. *Cálculo de la integral exponencial sobre la superficie de la esfera de radio h .*

$$\int_{\partial B_h} e^{i(\mathbf{x}'\boldsymbol{\omega})} ds(\mathbf{x}) = \left(\frac{2\pi h}{\|\boldsymbol{\omega}\|} \right)^{d/2} \|\boldsymbol{\omega}\| J_{\frac{d-2}{2}}(h\|\boldsymbol{\omega}\|).$$

Demostración. El cálculo de esta integral puede hacerse pasando a coordenadas hipersféricas, llegando así a una representación integral de la Función de Bessel de primera especie (Abramowitz y Stegun 1972, Ec. 9.1.20; Gradshteyn y Ryzhik 2007, Sec. 8.4):

$$J_\nu(z) = \frac{(z/2)^\nu}{\sqrt{\pi}\Gamma(\nu + 1/2)} \int_0^\pi \cos(z \cos \theta) (\sin \theta)^{2\nu} d\theta.$$

Puesto que $\boldsymbol{\omega}$ es fijo, siempre se puede poner el sistema de referencia de modo que $\mathbf{x}'\boldsymbol{\omega} = r\|\boldsymbol{\omega}\| \cos \phi_1$, donde r es el módulo de \mathbf{x} y ϕ_1 es el primero de los ángulos hipersféricos. El dominio de integración lo descompondremos en

dos partes $\Phi = [0, \pi]^{d-2} \times [0, 2\pi) = [0, \pi] \times \Phi_{-1}$ para integrar separadamente la coordenada ϕ_1 . Así,

$$\int_0^\pi \int_{\Phi_{-1}} e^{ir\|\boldsymbol{\omega}\| \cos \phi_1} r^{d-1} (\sin \phi_1)^{d-2} (\sin \phi_2)^{d-3} \cdots \sin \phi_{d-2} d\phi_1 \cdots d\phi_{d-1} =$$

$$r^{d-1} \left[\int_0^\pi (\sin \phi_1)^{d-2} e^{ir\|\boldsymbol{\omega}\| \cos \phi_1} d\phi_1 \right] \left[\int_{\Phi_{-1}} (\sin \phi_2)^{d-3} \cdots \sin \phi_{d-2} d\phi_2 \cdots d\phi_{d-1} \right].$$

La primera integral la descomponemos en sus partes real e imaginaria, mientras que la segunda es la superficie de la esfera unitaria en dimensión $d-1$, y por lo calculado en el Lema anterior vale $2\pi^{\frac{d-1}{2}}/\Gamma(\frac{d-1}{2})$. Se tiene entonces,

$$\frac{2\pi^{\frac{d-1}{2}} r^{d-1}}{\Gamma(\frac{d-1}{2})} \left[\int_0^\pi (\sin \phi_1)^{d-2} \cos(r\|\boldsymbol{\omega}\| \cos \phi_1) d\phi_1 + \right.$$

$$\left. i \int_0^\pi (\sin \phi_1)^{d-2} \sin(r\|\boldsymbol{\omega}\| \cos \phi_1) d\phi_1 \right].$$

La componente real se puede expresar en términos de la función de Bessel, con $\nu = \frac{d-2}{2}$ evaluada en $r\|\boldsymbol{\omega}\|$, como

$$\frac{\sqrt{\pi} \Gamma(\frac{d-2}{2} + \frac{1}{2})}{(r\|\boldsymbol{\omega}\|/2)^{\frac{d-2}{2}}} J_{\frac{d-2}{2}}(r\|\boldsymbol{\omega}\|).$$

Por su parte, la componente imaginaria es nula. En efecto, si $f(\phi_1) = (\sin \phi_1)^{d-2} \sin(r\|\boldsymbol{\omega}\| \cos \phi_1)$,

$$f(\pi - \phi_1) = \underbrace{(\sin(\pi - \phi_1))^{d-2}}_{\sin \phi_1} \sin(r\|\boldsymbol{\omega}\| \underbrace{\cos(\pi - \phi_1)}_{-\cos \phi_1}) = -f(\phi_1).$$

De modo que, las integrales sobre cada mitad del intervalo se anulan mutuamente, lo que puede verse con el cambio de variable $\psi = \pi - \phi_1$.

$$\int_{\frac{\pi}{2}}^{\pi} f(\phi_1) d\phi_1 = - \int_{\frac{\pi}{2}}^0 f(\pi - \psi) d\psi = - \int_0^{\frac{\pi}{2}} f(\psi) d\psi.$$

Reagrupando y cancelando factores se obtiene el resultado. \square

2.3.1 La función de correlación Exponencial en el espacio Euclídeo

La función de correlación exponencial tiene la siguiente forma

$$C(h) = e^{-\phi h},$$

donde $\phi, h > 0$. h representa la distancia (positiva) entre localizaciones y ϕ es el parámetro de *decaimiento*, que determina el rango efectivo R_{eff} .¹

$$C(R_{\text{eff}}) = 0,05 \times C(0) = \frac{1}{20} \iff -\phi R_{\text{eff}} = -\log 20 \iff R_{\text{eff}} \approx \frac{3}{\phi}.$$

Esta función es definida positiva en el espacio Euclídeo \mathbb{R}^d , $n \geq 1$ y por el Teorema de Bochner (Bochner, 1955, 1959) tendrá una representación espectral

$$C(\mathbf{x}) = \int_{\mathbb{R}^d} e^{i\langle \mathbf{x}, \boldsymbol{\omega} \rangle} dF(\boldsymbol{\omega}), \quad \forall \mathbf{x} \in \mathbb{R}^d, \quad (2.8)$$

donde \langle, \rangle denota el producto Euclídeo y F es una medida finita y no negativa en \mathbb{R}^d . Como es una función isotrópica, la dependencia respecto del vector \mathbf{x} es sólo a través de su módulo $h = \|\mathbf{x}\| = \sqrt{x_1^2 + \dots + x_d^2}$.

¹Suelen utilizarse otras parametrizaciones en las cuales ϕ cambia de significado. Por ejemplo, si $C(h) = e^{-h/\phi}$, entonces $R_{\text{eff}} \approx 3\phi$ y se suele denominar a ϕ como *rango*, aunque sólo sea proporcional al rango efectivo. Y si $C(h) = e^{-3h/\phi}$, entonces ϕ se corresponde prácticamente con el rango efectivo.

A continuación se calcula la medida espectral F , mediante transformación inversa de Fourier. Asumiremos que F es absolutamente continua, de modo que $dF(\boldsymbol{\omega}) = f(\boldsymbol{\omega})d\boldsymbol{\omega}$. Es una función que depende de la dimensión del espacio Euclídeo subyacente, por lo que calculamos en primer lugar el resultado en una dimensión y a continuación su generalización.

Dimensión 1.

$$\begin{aligned}
 f(\omega) &= \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\omega x} C(|x|) dx = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\omega x - \phi|x|} dx \\
 &= \frac{1}{2\pi} \left[\int_{-\infty}^0 e^{(\phi - i\omega)x} dx + \int_0^{\infty} e^{-(\phi + i\omega)x} dx \right] \\
 &= \frac{1}{2\pi} \left[\frac{1}{\phi - i\omega} [e^{(\phi - i\omega)x}]_{-\infty}^0 - \frac{1}{\phi + i\omega} [e^{-(\phi + i\omega)x}]_0^{\infty} \right] \\
 &= \frac{1}{2\pi} \left[\frac{1}{\phi - i\omega} + \frac{1}{\phi + i\omega} \right] \\
 &= \frac{\phi}{\pi(\phi^2 + \omega^2)}.
 \end{aligned}$$

Así, la representación espectral (2.8) para la función exponencial en dimensión 1 es:²

$$e^{-\phi|x|} = \frac{1}{\pi} \int_{\mathbb{R}} e^{ix\omega} \frac{\phi}{\phi^2 + \omega^2} d\omega. \quad (2.9)$$

Dimensión d .

$$\begin{aligned}
 f(\boldsymbol{\omega}) &= (2\pi)^{-d} \int \dots \int_{\mathbb{R}^d} e^{-i\langle \boldsymbol{\omega}, \mathbf{x} \rangle} C(\|\mathbf{x}\|) dx_1 \dots dx_d \\
 &= (2\pi)^{-d} \int \dots \int_{\mathbb{R}^d} e^{-i(\omega_1 x_1 + \dots + \omega_d x_d)} e^{-\phi \sqrt{x_1^2 + \dots + x_d^2}} dx_1 \dots dx_d.
 \end{aligned}$$

Sin pérdida de generalidad se puede asumir que $\boldsymbol{\omega} = \|\boldsymbol{\omega}\|e_1$, donde $e_1 = (1, 0, \dots, 0)$, pues dado $\boldsymbol{\omega}$, se puede hacer una rotación del sistema de coordenadas, de modo que el primer eje coordenado coincida con la dirección y el sentido de

² A modo de comprobación, la integral figura en [Gradshteyn y Ryzhik, 2007](#), 3.354(5).

ω . Como una rotación es un movimiento ortogonal en el espacio Euclídeo, el jacobiano de la transformación es igual a 1. Además, el módulo del vector \mathbf{x} no cambia y el producto escalar $\langle \omega, \mathbf{x} \rangle$ se reduce a $\|\omega\|x_1$.

Cambiando a coordenadas esféricas d -dimensionales se tiene $x_1 = \rho \cos \theta_1$ y jacobiano igual a $\rho^{d-1} \sin^{d-2} \theta_1 \cdots \sin \theta_{d-2}$. Por tanto, siendo $\Theta = [0, \pi)^{d-2} \times [0, 2\pi)$ la región de integración de las $d-1$ variables angulares, se pueden integrar todas excepto la primera, dando como resultado la superficie esférica unidad en \mathbb{R}^{d-1} .

$$\begin{aligned} f(\omega) &= (2\pi)^{-d} \int_0^\infty \int_\Theta e^{-i\|\omega\|\rho \cos \theta_1} e^{-\phi\rho} \rho^{d-1} \sin^{d-2} \theta_1 \cdots \sin \theta_{d-2} d\rho \cdots d\theta_{d-1} \\ &= (2\pi)^{-d} \cdot \frac{2\pi^{\frac{d-1}{2}}}{\Gamma(\frac{d-1}{2})} \cdot \int_0^\infty \rho^{d-1} e^{-\phi\rho} \int_0^\pi e^{-i\|\omega\|\rho \cos \theta_1} \sin^{d-2} \theta_1 d\theta_1 d\rho. \end{aligned}$$

La componente compleja de la integral interior se anula porque el integrando es antisimétrico respecto de $\theta = \frac{\pi}{2}$. La componente real se puede expresar en términos de una función de Bessel del primer tipo en lo que se conoce como la *Integral de Poisson* (Watson 1944, p. 24; Abramowitz y Stegun 1972, Ec. 9.1.20; Weisstein 2010):³

$$J_\nu(z) = \frac{(z/2)^\nu}{\Gamma(\frac{1}{2})\Gamma(\nu + \frac{1}{2})} \int_0^\pi \cos(z \cos \theta) \sin^{2\nu} \theta d\theta, \quad \nu \geq 0. \quad (2.10)$$

Entonces, para $\nu = \frac{d-2}{2}$ y $z = \rho\|\omega\|$ y teniendo en cuenta que $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, se tiene

$$f(\omega) = \frac{1}{(2\pi)^{\frac{d}{2}} \|\omega\|^{\frac{d-2}{2}}} \int_0^\infty e^{-\phi\rho} J_{\frac{d-2}{2}}(\|\omega\|\rho) \rho^{\frac{d}{2}} d\rho.$$

³ Estrictamente, en este punto tendría que distinguir el caso $d = 2$, ya que entonces hay una única variable angular y $\Theta = [0, 2\pi)$; entonces, la integral interior es entre 0 y 2π , que por la periodicidad de las funciones trigonométricas es justamente el doble que la integral entre 0 y π . Sin embargo esto se compensa con la superficie de la esfera 0-dimensional, que aunque surge de integrar el resto de las variables angulares (ninguna), la fórmula da como resultado precisamente 2.

La integral es un caso especial de la *Integral de Hankel* (Watson, 1944, p.386):

$$\int_0^\infty e^{-at} J_\nu(bt) t^{\nu+1} dt = \frac{2a \cdot (2b)^\nu \Gamma(\nu + \frac{3}{2})}{(a^2 + b^2)^{\nu + \frac{3}{2}} \Gamma(\frac{1}{2})}, \quad \phi > 0, \nu \geq 0.$$

Entonces, para $\nu = \frac{d-2}{2}$, $a = \phi$ y $b = \|\boldsymbol{\omega}\|$ se tiene

$$f(\boldsymbol{\omega}) = \frac{\Gamma(\frac{d+1}{2})}{\pi^{\frac{d+1}{2}}} \frac{\phi}{(\phi^2 + \|\boldsymbol{\omega}\|^2)^{\frac{d+1}{2}}}, \quad d \geq 1. \quad (2.11)$$

Así, la representación espectral (2.8) para la función exponencial en dimensión $d \geq 1$ es

$$e^{-\phi\|\boldsymbol{x}\|} = \frac{\Gamma(\frac{d+1}{2})}{\pi^{\frac{d+1}{2}}} \int \dots \int_{\mathbb{R}^d} e^{i\langle \boldsymbol{x}, \boldsymbol{\omega} \rangle} \frac{\phi}{(\phi^2 + \|\boldsymbol{\omega}\|^2)^{\frac{d+1}{2}}} d\boldsymbol{\omega}, \quad \phi \geq 0. \quad (2.12)$$

Notar que la Fórmula (2.12) se reduce a la contenida en la Ecuación (2.9) cuando $d = 1$. Para $d = 2$ se tiene, en cambio,

$$f_2(\boldsymbol{\omega}) = \frac{1}{2\pi} \cdot \frac{\phi}{(\phi^2 + \|\boldsymbol{\omega}\|^2)^{\frac{3}{2}}}.$$

Comprobación de la Fórmula (2.12)

Asumiendo sin pérdida de generalidad que $\boldsymbol{x} = \|\boldsymbol{x}\| \mathbf{e}_1$ y pasando a coordenadas esféricas d -dimensionales:

$$\begin{aligned} \omega_1 &= \rho \cos \theta_1 \\ \omega_2 &= \rho \sin \theta_1 \cos \theta_2 \\ &\vdots \\ \omega_{d-1} &= \rho \sin \theta_1 \sin \theta_2 \cdots \sin \theta_{d-2} \cos \theta_{d-1} \\ \omega_d &= \rho \sin \theta_1 \sin \theta_2 \cdots \sin \theta_{d-2} \sin \theta_{d-1} \end{aligned}$$

donde $\theta_i \in [0, \pi)$, $i = 1, \dots, d-2$ y $\theta_{d-1} \in [0, 2\pi)$. Entonces se producen los siguientes cambios

$$\begin{aligned}\langle \mathbf{x}, \boldsymbol{\omega} \rangle &= \|\mathbf{x}\| \rho \cos \theta_1, \\ \|\boldsymbol{\omega}\|^2 &= \rho^2, \\ d\boldsymbol{\omega} &= \rho^{d-1} \sin \theta_1^{d-2} \cdots \sin \theta_{d-3}^2 \sin \theta_{d-2} d\rho d\theta_1 \cdots d\theta_{d-1}.\end{aligned}$$

Así, llamando $\Theta = [0, \pi)^{d-2} \times [0, 2\pi)$ al dominio de las variables angulares, el integrando sólo es función de ρ y de θ_1 , permitiendo integrar el resto de las variables angulares en lo que resulta la superficie de la esfera unidad en el espacio de dimensión $d-1$:

$$\begin{aligned}\frac{\Gamma(\frac{d+1}{2})}{\pi^{\frac{d+1}{2}}} \int \cdots \int_{\mathbb{R}^d} e^{i\langle \mathbf{x}, \boldsymbol{\omega} \rangle} \frac{\phi}{(\phi^2 + \|\boldsymbol{\omega}\|^2)^{\frac{d+1}{2}}} d\boldsymbol{\omega} &= \\ &= \frac{\Gamma(\frac{d+1}{2})}{\pi^{\frac{d+1}{2}}} \int_0^\infty \int \cdots \int_{\Theta} e^{i\|\mathbf{x}\| \cos \theta_1} \frac{\phi}{(\phi^2 + \|\boldsymbol{\omega}\|^2)^{\frac{d+1}{2}}} \\ &\quad \rho^{d-1} \sin \theta_1^{d-2} \cdots \sin \theta_{d-3}^2 \sin \theta_{d-2} d\rho d\theta_1 \cdots d\theta_{d-1} \\ &= \frac{\Gamma(\frac{d+1}{2})}{\pi^{\frac{d+1}{2}}} \int_0^\infty \frac{\phi \rho^{d-1}}{(\phi^2 + \|\boldsymbol{\omega}\|^2)^{\frac{d+1}{2}}} \int_0^\pi e^{i\|\mathbf{x}\| \cos \theta_1} \sin \theta_1^{d-2} d\theta_1 \quad d\rho \cdot \frac{2\pi^{\frac{d-1}{2}}}{\Gamma(\frac{d-1}{2})}.\end{aligned}$$

Igual que en el cálculo directo, la integral interior se reduce a la *Integral de Poisson* (2.10), resultando entonces,

$$= \frac{2^{d/2} \Gamma(\frac{d+1}{2})}{\pi^{1/2} \|\mathbf{x}\|^{\frac{d-2}{2}}} \int_0^\infty \rho^{d/2} J_{\frac{d-2}{2}}(\|\mathbf{x}\| \rho) \frac{\phi}{(\phi^2 + \rho^2)^{\frac{d+1}{2}}} d\rho.$$

Esta última expresión es calculada por el programa de cálculo simbólico *Mathematica* (Wolfram Research, Inc., 2008) como $e^{-\phi\|\mathbf{x}\|}$. Además, la integral figura en *Gradshteyn y Ryzhik* (2007, 6.565(3)). De este modo se confirma doblemente la validez de la Fórmula (2.12).

2.4 Predicción geoestadística: Kriging

Una vez que se ha ajustado un modelo a los datos, interesa predecir en localizaciones no observadas, generalmente a los efectos de construir un mapa completo del fenómeno en la región de interés. Hay varios enfoques para este problema, pero el método de predicción espacial más ampliamente utilizado se conoce como *Kriging*. La mayoría de los métodos utilizan una media ponderada de los valores muestrales para generar la predicción; aquellas observaciones próximas a la localización de predicción tienen así un peso mayor que las más alejadas. El Kriging determina estos pesos basándose en la función de semivariograma.

Su popularidad se debe en gran medida a algunas propiedades estadísticas que hacen atractivo el predictor Kriging. Además de ser el mejor predictor lineal insesgado (BLUP, por sus siglas en inglés) en términos del error cuadrático, es notablemente robusto ante violaciones de las hipótesis del modelo (Cressie y Zimmerman, 1992) y proporciona una estimación del error estándar de predicción en cada punto.

Existen varios tipos de Kriging: Ordinario, Simple, Universal, en Bloques y otros. Aquí desarrollamos el Kriging Universal, que es más general que el ordinario y el simple y es el utilizado en el trabajo.

El método asume que el vector de observaciones $\mathbf{Z} = (Z(\mathbf{s}_1), \dots, Z(\mathbf{s}_n))$ viene generado por un proceso Gaussiano intrínsecamente estacionario e isotrópico con función de media $\mu(\mathbf{s}) = \beta_0 + \beta_1 f_1(\mathbf{s}) + \dots + \beta_p f_p(\mathbf{s})$ y un semivariograma conocido $\gamma(r)$, donde las $f_i(\cdot)$ son funciones de la localización espacial \mathbf{s} o variables explicativas asociadas a las localizaciones.

Desde un punto de vista estadístico, el predictor kriging en una localización dada \mathbf{s}_0 se define naturalmente como la esperanza condicional del valor del proceso en \mathbf{s}_0 con respecto a su valor en los puntos observados

$$\hat{Z}(\mathbf{s}_0) = \mathbb{E}[Z(\mathbf{s}_0)|\mathbf{Z}].$$

Anotando donde $\boldsymbol{\gamma} = (\gamma(\|\mathbf{s}_1 - \mathbf{s}_0\|), \dots, \gamma(\|\mathbf{s}_n - \mathbf{s}_0\|))'$ y $\boldsymbol{\Gamma} = (\gamma(\|\mathbf{s}_i - \mathbf{s}_j\|))$, el modelo de semivariograma determina la matriz de covarianzas

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma^2 & \boldsymbol{\gamma}' \\ \boldsymbol{\gamma} & \boldsymbol{\Gamma} \end{pmatrix}$$

del vector normal multivariante $(Z(\mathbf{s}_0), \mathbf{Z})$. Y utilizando una propiedad estándar de la distribución normal multivariante, la distribución condicional de $Z(\mathbf{s}_0)$ dadas las observaciones \mathbf{Z} es también normal multivariante con media

$$\mathbb{E}[Z(\mathbf{s}_0)|\mathbf{Z}] = \mu(\mathbf{s}_0) + \boldsymbol{\gamma}'\boldsymbol{\Gamma}^{-1}(\mathbf{Z} - \mu(\mathbf{Z}))$$

y varianza

$$\mathbb{V}[Z(\mathbf{s}_0)|\mathbf{Z}] = \sigma^2 - \boldsymbol{\gamma}'\boldsymbol{\Gamma}^{-1}\boldsymbol{\gamma}.$$

Este enfoque estadístico, sistematizado por [Matheron \(1963\)](#), exige un modelo subyacente válido. Esto es, se asume la existencia de un proceso Gaussiano subyacente con un modelo de semivariograma condicionalmente negativo que produzca matrices de covarianzas $\boldsymbol{\Sigma}$ definidas positivas, para cualquier posible elección de puntos de predicción en la región de trabajo. En la metodología desarrollada en el capítulo siguiente, estas condiciones se pondrán en cuestión, y por tanto el predictor Kriging y la correspondiente varianza de predicción pierden su fundamento teórico.

Sin embargo, el predictor kriging tiene su origen en el campo de la Ingeniería, donde se deduce mediante un cálculo de optimización, minimizando el error de predicción cuadrático medio. Cuando el modelo estadístico subyacente está bien definido, es sencillo ver que los dos enfoques coinciden [Diggle y Ribeiro \(2007, Teo. 6.1\)](#). Sin embargo, el enfoque *ingenieril* no asume la existencia de un modelo estadístico subyacente, y por tanto los resultados siguen siendo válidos siempre que la matriz de covarianzas observada sea definida positiva. Este punto de vista, desarrollado a continuación, proporciona la interpretación adecuada de la metodología desarrollada en el capítulo siguiente.

El predictor Kriging en una localización dada \mathbf{s}_0 se escribe como combinación lineal de los valores observados en las localizaciones $\mathbf{s}_1, \dots, \mathbf{s}_n$:

$$\hat{Z}(\mathbf{s}_0) = \sum_{i=1}^n \lambda_i Z(\mathbf{s}_i),$$

donde los $\lambda_1, \dots, \lambda_n$ se eligen para minimizar el error de predicción cuadrático medio

$$\mathbb{E} \left[(\hat{Z}(\mathbf{s}_0) - Z(\mathbf{s}_0))^2 \right],$$

sujetos a la condición $\sum_{i=1}^n \lambda_i = 1$ que asegura que el predictor sea insesgado, gracias a la estacionariedad de la media del proceso:

$$\mathbb{E} \left[\hat{Z}(\mathbf{s}_0) \right] = \mathbb{E} \left[\sum_{i=1}^n \lambda_i Z(\mathbf{s}_i) \right] = \sum_{i=1}^n \lambda_i \mathbb{E} [Z(\mathbf{s}_i)] \stackrel{i)}{=} \sum_{i=1}^n \lambda_i \mu = \mathbb{E} [Z(\mathbf{s}_0)].$$

Este problema de optimización conduce a un sistema de ecuaciones con restricción. En el Apéndice A proporcionamos una deducción detallada de este sistema en los casos Simple y Ordinario. El caso Universal se deduce del Ordinario sustituyendo la media constante por una estimación que depende de la matriz de diseño del modelo lineal. De modo que la solución del sistema resulta (Cressie, 1993; Diggle y Ribeiro, 2007):

$$\hat{Z}(\mathbf{s}_0) = \left[\boldsymbol{\gamma} + \mathbf{X}(\mathbf{X}'\boldsymbol{\Gamma}^{-1}\mathbf{X})^{-1}(\mathbf{x} - \mathbf{X}'\boldsymbol{\Gamma}^{-1}\boldsymbol{\gamma}) \right]' \boldsymbol{\Gamma}^{-1} \mathbf{Z},$$

con varianza de predicción:

$$\sigma^2(\mathbf{s}_0) = \boldsymbol{\gamma}'\boldsymbol{\Gamma}^{-1}\boldsymbol{\gamma} - (\mathbf{1}'\boldsymbol{\Gamma}^{-1}\boldsymbol{\gamma} - 1)^2 / (\mathbf{1}'\boldsymbol{\Gamma}^{-1}\mathbf{1}),$$

donde $\boldsymbol{\gamma} = (\gamma(\|\mathbf{s}_1 - \mathbf{s}_0\|), \dots, \gamma(\|\mathbf{s}_n - \mathbf{s}_0\|))'$, $\mathbf{x} = (f_0(\mathbf{s}_0), f_1(\mathbf{s}_0), \dots, f_p(\mathbf{s}_0))'$, $\boldsymbol{\Gamma} = \left(\gamma(\|\mathbf{s}_i - \mathbf{s}_j\|) \right)$, y $\mathbf{X} = \left(f_{j-1}(\mathbf{s}_i) \right)$, siendo $f_0(\mathbf{s}) = 1 \quad \forall \mathbf{s}$.

La varianza de predicción proporciona una medida de la capacidad predictiva del modelo. Sin embargo, esta medida es válida bajo las hipótesis del modelo

(normalidad, estacionariedad, etc.) y suponiendo conocidos los parámetros del variograma. Por este motivo, la predicción Kriging suele evaluarse mediante métodos de *validación cruzada*. Esto es, suprimir la i -ésima de las observaciones y predecir su valor utilizando el modelo ajustado con las restantes. Se puede calcular así el *error medio de predicción* como la media aritmética de los errores cometidos en la predicción de cada una de las observaciones y utilizar este valor como medida de la capacidad predictiva del modelo.

Cuando la predicción se realiza bajo el paradigma Bayesiano, la incertidumbre en los parámetros del variograma se traslada adecuadamente a la varianza de predicción, y la validación cruzada sólo es necesario hacerla cuando existen serias dudas sobre las hipótesis de partida. Especialmente sobre la estacionariedad.

2.5 Conclusiones

En este capítulo hemos introducido los métodos y conceptos geoestadísticos clásicos y se proporciona el conjunto básico de herramientas necesario para el desarrollo de la Tesis. No debe considerarse un tratado exhaustivo sobre la Geoestadística. Se han dejado al margen múltiples aspectos importantes, por considerarlos prescindibles para el tema que nos ocupa.

Se hace especial énfasis en la Teoría Espectral, demostrando con detalle los teoremas de Bochner y Schoenberg, que serán centrales en el desarrollo de la Tesis. También se calcula la representación espectral de la función exponencial en el espacio Euclídeo de una o varias dimensiones, que será una referencia importante en capítulos posteriores.

Capítulo 3

Geoestadística basada en el coste

El proyecto que dio lugar a esta Tesis consistió en la exploración de una solución metodológica aplicada para la elaboración de mapas acústicos, particularmente en entornos urbanos.

Los métodos habitualmente utilizados consisten en la simulación por ordenador de la generación y difusión del ruido a partir de parámetros conocidos o estimados acerca de las fuentes de ruido, especialmente el tráfico vehicular, ferroviario, aeronáutico y la actividad industrial. En el mejor de los casos se utilizan mediciones de campo para calibrar estos parámetros.

Sin embargo, la valoración del error de predicción es comúnmente omitida, o estimada bajo condiciones muy ideales. Consideramos que la Geoestadística es una herramienta oportuna para esta valoración de la incertidumbre, que por otra parte es importante para garantizar con cierta certeza una exposición mínima de la población a niveles insalubres de ruido.

Por otra parte, la predicción en entornos urbanos introduce desafíos importantes para la metodología geoestadística. Este aspecto, que puede abstraerse

como un problema de predicción geoestadística en regiones heterogéneas, es el que ha despertado nuestro interés y ha motivado los capítulos posteriores.

En este capítulo presentamos el proyecto que utilizamos como banco de pruebas y detallamos la metodología práctica que desarrollamos para resolver el problema de la heterogeneidad de las áreas urbanas.

3.1 Proyecto piloto

3.1.1 Mapas acústicos mediante técnicas geoestadísticas

La *contaminación acústica* en áreas urbanas es un motivo de creciente preocupación por parte de las autoridades ([World Health Organization, 1999](#)). Los mapas acústicos son las herramientas de diagnóstico utilizadas para la planificación de acciones preventivas y correctoras. Éstos representan, para cada localización, el nivel medio de ruido registrado en un intervalo de tiempo determinado, en una escala apropiada.

La medición del nivel de ruido no es una tarea trivial. Existen muchas restricciones a tener en cuenta, y es una operación que consume no menos de 15 o 30 minutos de un operario calificado. En resumen, las observaciones de campo son costosas.

Esto ha impulsado otra línea de trabajo basada en la simulación de modelos deterministas de difusión del ruido, que hacen uso de un modelo digital de la ciudad, en conjunto con una gran cantidad de información relativa a las características del tráfico en cada una de las vías. Sin embargo, la incertidumbre y la variabilidad propia de los parámetros y de las simplificaciones del modelo se propagan a lo largo de miles de iteraciones de una forma desconocida y descontrolada.

En general, se pueden clasificar los métodos de elaboración de mapas acústicos en *observacionales* o *de cálculo* según se basen en mediciones de campo, o en simulación por ordenador, respectivamente. Sin embargo, lo más habitual es combinar ambas estrategias en los llamados *métodos híbridos*.

La *Directiva 2002/49/CE* del Parlamento Europeo y del Consejo, de 25 de junio de 2002, *sobre evaluación y gestión del ruido ambiental* ([Parlamento y Consejo Europeos, 2002](#)), en su anexo segundo, describe métodos provisionales para la determinación de los indicadores sonoros estándar definidos en la norma *ISO 1996—2: 1987*, tanto por la vía de la medición como del cálculo. Estos indicadores, o derivados de los mismos (como regiones que superan determinado umbral, número de personas, viviendas, colegios u hospitales afectados, etc.) son los que se representan gráficamente en los mapas acústicos resultantes.

Métodos observacionales

Las mediciones se realizan con un dispositivo llamado *sonómetro integrador* en diferentes puntos, típicamente en un arreglo aproximadamente regular cubriendo la región de interés (ver Fig. 3.1), que suele afinarse en zonas de especial aglomeración de personas y/o tráfico.

La medición con estos dispositivos conlleva un protocolo riguroso para asegurar la calidad y la comparabilidad de las observaciones. Hay que hacerlo a una altura determinada del suelo, a cierta distancia de las fachadas, durante un cierto lapso de tiempo en el cual no pueden haber eventos anómalos, en determinadas condiciones meteorológicas, y conforme a una lista de restricciones de procedimiento detalladas en la Directiva Europea ya mencionada y en la normas *ISO 1996—2: 1987* e *ISO 1996—1: 1982*. En cada medición se registran los niveles promedio de varios indicadores sonoros, expresados en *decibelios* dB(A).

Las mediciones así obtenidas son muy fiables, ya que en la práctica el error de medida es despreciable. En contrapartida, el valor obtenido es muy variable con la franja horaria, día de la semana, época del año, etc. Por otra parte, obtener cada observación resulta costoso, ya que los sonómetros son dispositivos sofisticados y caros, cuyo manejo requiere de personal especializado y porque cada medición consume un tiempo considerable, a veces ha de repetirse, y no se puede realizar en cualquier momento ni bajo cualquier circunstancia.

Una vez realizadas las mediciones no existe un procedimiento estandarizado ni oficial para interpolar las observaciones. Una manera práctica, aunque grosera,

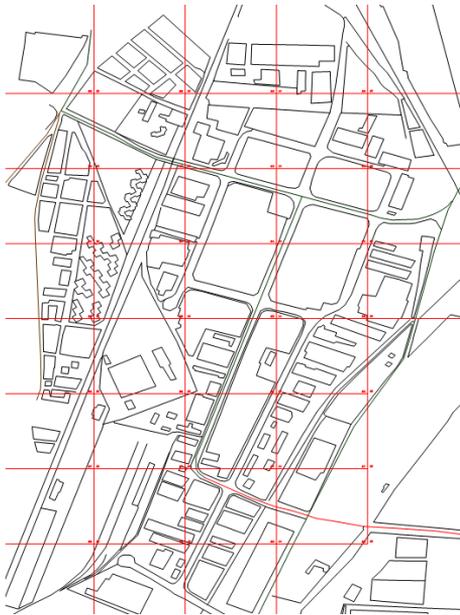


Figura 3.1. Disposición de los puntos de medición en un arreglo regular.

que se ha utilizado es asignar a cada celda de la cuadrícula regular el valor medio de las mediciones hechas en las esquinas. Por supuesto no hay garantía alguna de que toda la zona cubierta por la celda sea mínimamente homogénea, ni se puede controlar el error que se comete.

Métodos de cálculo

Este tipo de métodos, en cambio, se aplican de forma completamente virtual en un ordenador, evitando la observación directa. En contrapartida requieren un modelo digital de la ciudad lo más detallado posible (ver Fig. 3.2) y abundante información sobre características del entorno y de las fuentes sonoras.

Para ejemplificar, la información necesaria incluye la topografía de la zona a evaluar, la climatología, todos los obstáculos (edificios, árboles, casetas, puentes, carreteras, calles, plazas, estaciones de autobuses y trenes, etc.) y parámetros como el número de vehículos por unidad de tiempo que circulan por cada una de las calles de la zona según el tipo (camiones, autobuses y vehículos pesados, turismos o motocicletas), velocidad media de la vía, pendiente, tipo de asfalto, etc.

Con software especializado se simula la propagación del ruido provocado por cada fuente sonora (únicamente de los siguientes tipos, que son los más relevantes: tráfico rodado, trenes, aeronaves e industrias), teniendo en cuenta los fenómenos acústicos de atenuación, absorción aérea, absorción en contornos, reflexión, difracción y refracción (ver Fig. 3.3). Muchos de estos efectos dependen de la frecuencia, por lo que la simulación se realiza para distintas bandas de frecuencia. Así se calcula la contribución de cada fuente sonora al ruido ambiente en cada punto de la región (en la práctica, en una cuadrícula suficientemente fina de puntos).

A pesar de que estos métodos requieren de muchas horas de cálculo intensivo, el proceso en sí no tiene, en principio, un coste directo asociado. Aunque se podría considerar como coste indirecto el asociado a la disponibilidad de tanta información detallada sobre las características de la ciudad.



Figura 3.2. Los métodos de cálculo requieren un modelo digital detallado de la ciudad.

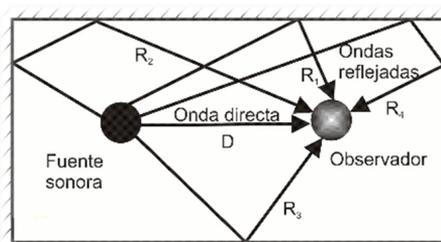


Figura 3.3. La propagación del ruido se simula teniendo en cuenta los fenómenos acústicos.

Una ventaja fundamental de estos métodos es que permiten hacer simulación sobre actuaciones urbanísticas proyectadas, con el objetivo de predecir su impacto futuro. Algo inaccesible para los métodos observacionales.

Los procedimientos de cálculo de mapas acústicos están diseñados fundamentalmente para el ruido de carreteras, ferrocarriles, aeropuertos e industrias. La contaminación acústica urbana es mucho más y muy diferente a eso. Para valorarla es preciso incluir otras muchas fuentes, imposibles de ser, a día de hoy, modeladas por procedimientos de cálculo: el comportamiento ciudadano en la vía pública, la existencia de bocinas y sirenas, la concentración de actividades, la existencia de pasos de peatones, de paradas de autobuses, de semáforos, de túneles, etc., no pueden ser introducidos en estos programas y son factores, en muchos casos, determinantes de los niveles sonoros ambientales de una determinada zona.

Las estimaciones obtenidas por los métodos de cálculo dependen fuertemente de la calidad de los datos de entrada, en particular de los parámetros característicos de las fuentes sonoras. Es habitual por tanto, la realización de mediciones específicas que permitan ajustar estos parámetros. Estas mediciones son distintas de las realizadas en los métodos observacionales, ya que su finalidad es la evaluación del ruido *emitido* por la fuente.

La información sobre los parámetros contiene una variabilidad inherente, producto de la simplificación y parametrización. Por ejemplo, el número de vehículos por unidad de tiempo que circulan por una vía no es constante, la velocidad del tráfico es sólo una media, y así sucesivamente con todos los parámetros. También el modelo del entorno físico-geométrico tiene un cierto error, y finalmente el método de simulación utilizado también introduce algunas simplificaciones imprescindibles, como por ejemplo el hecho de que una fuente de ruido propaga el sonido a través de unos 500 rayos en una cantidad finita de ángulos posibles, que se pueden reflejar a lo sumo unas cuantas veces y que tienen efecto hasta una distancia máxima. Todas estas incertidumbres interactúan y se propagan a través de miles de iteraciones de una forma desconocida. Los resultados obtenidos son aquellos que resultan de suponer que todos los datos son exactos. Pese a ello, los vendedores de software especializado afirman que

los resultados tienen una incertidumbre de 2 o 3 dB(A), *hasta distancias medias y siempre que los datos sean suficientemente buenos.*

Métodos híbridos

Los métodos híbridos basan sus predicciones en el cálculo de mapas, pero utilizando algunas medidas específicas con el objetivo de ajustar posibles desviaciones de los parámetros de entrada. Por ejemplo, se utilizan algunas estaciones permanentes de medición de ruido localizadas estratégicamente para ser afectadas por una única fuente de ruido específica, permitiendo así actualizar sus correspondientes parámetros mediante cálculo inverso. También se suelen utilizar medidas puntuales para controlar el error de simulación.

Como hemos mencionado en el capítulo anterior, la Geoestadística proporciona un conjunto de herramientas específicamente diseñadas para fenómenos espaciales, en los cuales se pretende predecir una cierta cantidad sobre una región de interés en la cual se han realizado un conjunto de observaciones. El nivel de ruido es una magnitud medible y que se distribuye espacialmente de forma continua. Tiene una evidente *estructura espacial*; esto es, que los niveles de ruido en cualquier conjunto de localizaciones están correlacionados entre sí. Además, tiene sentido modelar esta estructura de correlaciones en función de la distancia, ya que el nivel de ruido en localizaciones próximas es esperable que sea similar, mientras que en localizaciones lejanas los niveles pueden ser completamente independientes. Cabe notar que representar en un mapa un nivel de ruido que no se ha observado, es efectivamente un problema de predicción y por tanto está sujeto a error.

Estas características hacen que la modelización geoestadística se adapte perfectamente al problema de la elaboración de mapas acústicos a partir de un conjunto de mediciones del nivel de ruido. Pero no sólo es que sea un método adecuado al problema en cuestión, sino que tiene dos propiedades que son especialmente atractivas, gracias a que el procedimiento se basa en un modelo estadístico subyacente:

1. puede incorporar información adicional en la forma de covariables;
2. se pueden obtener estimaciones del error de predicción en cada punto.

La incorporación de información adicional permite hacer predicciones más afinadas, sobre todo con menor error de predicción, puesto que el modelo descompone el nivel de ruido en una parte no espacial, explicada por las covariables, y en una parte espacial que queda más nítidamente identificable, desprovista de *ruido*. Por otra parte, la estimación del error de predicción, es una información muy importante que es frecuentemente infravalorada.

3.1.2 Presentación de los datos

A los efectos de su utilización como datos de prueba hemos realizado un conjunto de 52 mediciones del nivel de ruido en varios puntos del barrio de Malilla en la ciudad de Valencia. Un subconjunto de ellas fueron tomadas muy cerca las unas de las otras a lo largo de dos calles: una avenida y una calle con baja intensidad de tráfico, con el objetivo de facilitar la estimación del semivariograma a distancias cortas (ver Fig. 3.4).

El barrio elegido no es homogéneo en lo que al ruido se refiere. De hecho, presenta importantes contrastes, con grandes avenidas cerca de calles pequeñas y tranquilas. Un paso elevado al norte y una vía férrea al oeste constituyen elementos influyentes. Estas características proporcionan una gran diversidad de situaciones que enriquecen el estudio.

El ruido generado por el tráfico rodado es habitualmente considerado como la principal fuente de ruido urbano. Por tanto las diferencias en la densidad del tráfico explican gran parte de las diferencias en los niveles de ruido. Por este motivo se han clasificado las vías de acuerdo con la densidad de tráfico media estimada. Las avenidas se consideran vías de categoría 1, mientras que a las calles con densidad de tráfico media y baja se les ha asignado las categorías 2 y 3 respectivamente.



Figura 3.4. Localizaciones observadas (círculos rojos) y clasificación de vías según la densidad de tráfico. Las líneas rojas representan avenidas con una alta densidad de tráfico, mientras que las amarillas y verdes son calles con densidades de tráfico media y baja respectivamente.

Nota. Debemos dejar constancia de que no todas las mediciones fueron realizadas en las condiciones reglamentarias a las que hicimos referencia en la página 43, por tanto los mapas acústicos resultantes no son necesariamente representativos de la realidad y deben interpretarse únicamente en el marco de este estudio. Por este motivo no consideramos relevante presentar la tabla de valores observados.

3.1.3 Predicción geoestadística

Puesto que la densidad de tráfico juega un papel tan relevante, hemos considerado oportuno trabajar con los datos normalizados con respecto a la categoría de la vía correspondiente, de forma de poder apreciar y evaluar mejor el efecto espacial. Concretamente trabajamos con los residuos de un modelo lineal con la categoría de vía como factor; o lo que es lo mismo, a cada dato se le ha quitado la media de todas las observaciones que están en vías de su misma categoría.

Una vez que se tienen las coordenadas de las observaciones y los valores observados (en este caso, normalizados), es inmediato ajustar un modelo de semivariograma y realizar la predicción geoestadística. En este caso se ha utilizado

el modelo de semivariograma Exponencial, aunque los resultados de la predicción no son especialmente sensibles a la elección de la familia de semivariograma.

En la Figura 3.5a se muestra el mapa de predicción resultante. Hay que notar especialmente que, dados los datos, los edificios y el resto de la infraestructura urbana no juega ningún papel en la predicción ya que la estructura de correlaciones se determina exclusivamente a partir de las distancias entre las localizaciones. Consecuentemente, el mapa resultante se extiende sobre toda la región, incluso en aquellas áreas donde no tiene sentido predecir. Por supuesto, lo razonable es hacer las predicciones sólo donde tiene sentido, o equivalentemente, descartar las que no tienen sentido superponiendo en el mapa la *capa* de edificaciones, como se muestra en la Figura 3.5b.

Hay que tener en cuenta que éstos mapas representan la predicción con las observaciones normalizadas. Esto es, sin tener en cuenta el efecto (predominante) del tráfico. Para obtener un verdadero mapa de ruido por este método es necesario volver a incorporar los valores medios estimados correspondientes a la categoría de la vía en la que se encuentran.

Esto puede hacerse de varias maneras, puesto que los puntos de predicción que están en intersecciones o alejados de las vías no están claramente asociados a ninguno de las categorías de vía. Vamos a posponer esta discusión para la Sección 3.3.4, ya que adoptaremos otra estrategia menos grosera para tener en cuenta la densidad del tráfico. De momento, este análisis preliminar resulta suficiente para nuestros propósitos.

3.1.4 Problema específico y motivación de la solución.

Por experiencia sabemos que las edificaciones influyen significativamente en el nivel de ruido. Si nos refugiamos detrás de un muro, o damos vuelta a la esquina apartándonos de una avenida ruidosa el nivel de ruido disminuye ostensiblemente. Despreciarlas en el modelo, prediciendo igual que si no existieran como si fuera en campo abierto, parece una simplificación excesiva.



Figura 3.5. Predicción geoestadística estándar (a) en toda la región y (b) sólo donde la predicción tiene sentido.

En el fondo, el problema parece producirse exactamente en el punto en el que modelamos la correlación entre dos localizaciones como una función de la distancia (ver la definición de la estacionariedad débil, en la página 10). En la Figura 3.6 se muestran esquemáticamente tres parejas de puntos situadas a la misma distancia euclídea entre sí y por tanto con una misma correlación estadística, de acuerdo con la modelización estándar. Sin embargo resulta evidente que las tres situaciones son cualitativamente muy diferentes. En el caso (b) hay una barrera entre las localizaciones, que podemos pensar como la abstracción de un edificio en nuestro problema original. No cabe duda que la correlación en este caso debe ser menor que en el caso (a), en el cual no hay ningún obstáculo. De la misma manera, si el obstáculo resulta mucho más grande, como en (c), esperaríamos que la correlación fuera aún menor.

La conclusión evidente es que en un entorno urbano, que está repleto de obstáculos que actúan como barreras acústicas, *la correlación entre las localizaciones no puede ser una función que depende únicamente de la distancia que las separa*. La estacionariedad débil parece una hipótesis demasiado simplista en este caso.

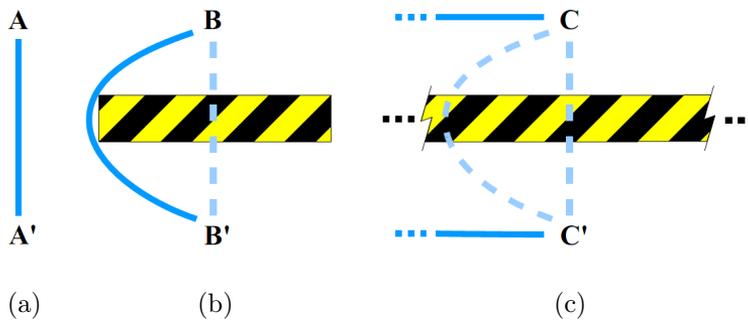


Figura 3.6. Esquema que muestra la inconveniencia de modelar la correlación como función de la distancia euclídea.

Una alternativa que surge naturalmente es asociar la correlación con la distancia mínima que hay que recorrer para ir de una localización a otra, sin cruzar ninguna barrera. O dicho de otra manera, la distancia mínima que recorre *el ruido* entre una localización y la otra, representada en la Figura 3.6 por las longitudes de las líneas continuas.

Varios investigadores han llegado a esta conclusión desde los trabajos de Little et al. (1997) y Rathbun (1998) en el campo del análisis geoestadístico en estuarios, donde han considerado más natural utilizar *distancias acuáticas*. Los siguieron Curriero (1996) y Krivoruchko y Gribov (2002). Curriero (2006) mostró que muchos de los modelos de covarianzas paramétricos tradicionales no son válidos con distancias no euclídeas. Por tanto, estas distancias no pueden utilizarse sin una prueba de validez del modelo. Otros autores como Ver-Hoef et al. (2004), Higdon (1998), Løland y Høst (2003) y, más recientemente, Okabe et al. (2009) han explorado diferentes estrategias como *Moving Window Kernels* o *Multidimensional Scaling*.

La utilización de este tipo de distancias no euclídeas tiene tres características importantes:

- Mantiene la idea intuitiva de que localizaciones próximas deben estar más correlacionadas que localizaciones distantes (Ley de Tobler, p. 8).

- Tiene en cuenta el entorno y lo hace de una manera razonable. En el esquema de la Figura 3.6, la pareja (c) está mucho más *alejada* que la (b) y ésta a su vez que la (a), lo que se traduce en correlaciones respectivamente menores.
- Es un concepto que generaliza al enfoque estándar y por tanto lo contiene como caso particular: cuando no hay barreras, la distancia mínima coincide con la euclídea.

El desafío es pues, desarrollar una metodología práctica para aplicar las técnicas geoestadísticas a partir de una función de correlación que dependa, no de la distancia euclídea, sino de esta generalización.

3.2 Distancia basada en el coste

3.2.1 Definición y propiedades

Aunque no es una hipótesis explícita de los métodos geoestadísticos, modelar la correlación como una función de la distancia euclídea implica la asunción de que la región de trabajo es homogénea y sin restricciones.

Se ha visto cómo las infraestructuras urbanas constituyen barreras acústicas, restricciones, que violan esta homogeneidad. Hay otras situaciones intermedias, en las que, sin haber barreras, la región presenta heterogeneidades, o irregularidades que producen el mismo efecto: hacer poco verosímil el modelo de correlación como función de la distancia euclídea.

Por ejemplo, un hongo se propagará fácilmente sobre áreas fértiles, cálidas y protegidas de un campo, pero es más difícil que lo haga en porciones rocosas y expuestas. Esta heterogeneidad puede modelarse con una *superficie de coste* que represente cuán difícil es atravesar una región dada para el fenómeno en cuestión, para entonces asociar la correlación entre localizaciones con la ruta de mínimo coste entre ellas. Formalmente,

Definición 3.1 (Superficie de coste). Una función f definida sobre la región de interés, con valores en los números reales positivos, de modo que su valor en una localización dada es interpretado como la *densidad de coste* en ese punto. Esto es, que el coste de realizar un desplazamiento infinitesimal ds en el punto s tiene un coste $f(s)ds$.

Esta es la herramienta que permite modelar cualquier factor que afecte la correlación. En particular, en esta superficie se sintetizan la distancia Euclídea y la configuración del entorno.

La superficie de coste no es necesariamente continua, ni siquiera acotada. Por ejemplo, las barreras pueden modelarse como regiones con coste infinito. Teóricamente la superficie de coste debería estar definida sobre todo el plano, ya que la ruta de mínimo coste debe buscarse entre todas las rutas posibles. En la práctica, sin embargo, es suficiente definirla sobre una región que cubra todas las localizaciones de interés, siempre que se pueda argumentar que todas las rutas de mínimo coste deben quedar contenidas dentro de dicha región.

Así, cualquier ruta entre dos localizaciones tiene un coste asociado.

Definición 3.2 (Coste de una ruta). Dada una superficie de coste, toda ruta contenida en la región de trabajo D tiene un coste asociado que se calcula integrando la superficie de coste a lo largo de ella.

Formalmente, si $\alpha : [0, 1] \rightarrow D \in C^1$ con $\alpha(0) = s_1$; $\alpha(1) = s_2$ es una ruta o trayectoria entre dos puntos $s_1, s_2 \in D$, el coste de la trayectoria α es

$$\int_0^1 f(\alpha(t)) \alpha'(t) dt.$$

Notar que una trayectoria debe ser naturalmente continua y diferenciable salvo, posiblemente, un número finito de puntos de *quiebre*, en cuyo caso la integral ha de entenderse *a trozos*.

Definición 3.3 (Distancia basada en el coste). Dada una superficie de coste, la distancia basada en el coste entre dos localizaciones se define como el coste de una ruta de mínimo coste entre ellas.

Una propiedad importante de la distancia basada en el coste es que es efectivamente una métrica, una distancia en el sentido matemático del término. Es la primera condición para realizar análisis geoestadísticos con distancias no Euclídeas (Waller y Gotway, 2004).

Propiedad 3.4 (La distancia basada en el coste es una métrica). *Dada una superficie de coste f definida en una región $D \subseteq \mathbb{R}^2$ y denotando por $\mathfrak{d}(\mathbf{x}, \mathbf{y})$ a la distancia basada en el coste entre los puntos $\mathbf{x}, \mathbf{y} \in D$ con respecto a la superficie f , se verifican los axiomas de una métrica: $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in D$*

$$i) \mathfrak{d}(\mathbf{x}, \mathbf{y}) \geq 0, \text{ y } \mathfrak{d}(\mathbf{x}, \mathbf{y}) = 0 \iff \mathbf{x} = \mathbf{y} \quad (\text{No negatividad})$$

$$ii) \mathfrak{d}(\mathbf{x}, \mathbf{y}) = \mathfrak{d}(\mathbf{y}, \mathbf{x}) \quad (\text{Simetría})$$

$$iii) \mathfrak{d}(\mathbf{x}, \mathbf{y}) \leq \mathfrak{d}(\mathbf{x}, \mathbf{z}) + \mathfrak{d}(\mathbf{z}, \mathbf{y}) \quad (\text{Desigualdad Triangular})$$

Demostración. *i)* La integral de una función positiva a lo largo de una curva es no negativa y es 0 si y sólo si la curva tiene longitud nula; esto es, si el punto de partida y llegada son el mismo, o sea que se está calculando la distancia basada en el coste de un punto consigo mismo. *ii)* El coste de una ruta no depende del sentido en que se recorra. *iii)* Si no fuera así, la ruta entre \mathbf{x} y \mathbf{y} que se obtiene uniendo las rutas de mínimo coste entre \mathbf{x} y \mathbf{z} y entre \mathbf{z} y \mathbf{y} , resulta ser, en virtud de la aditividad de la integral, una ruta de coste menor que la ruta de mínimo coste. Absurdo. \square

Finalmente, las condiciones clásicas de homogeneidad de la región de interés son un caso particular en el cual la superficie de coste es una superficie constante con valor 1 en todos los puntos. Consecuentemente, la ruta de mínimo coste entre dos localizaciones cualquiera es simplemente la *línea recta* que las conecta y por tanto la distancia basada en el coste se reduce a la distancia euclídea. La situación más general con barreras es otro caso particular donde la superficie de coste toma el valor 1 en todos los sitios, excepto en las propias barreras donde toma un valor infinito. De este modo, la distancia basada en el coste coincide con la mínima distancia que se ha de recorrer para conectar las dos localizaciones

sin cruzar ninguna barrera, coincidiendo con la adaptación intuitiva desarrollada en la Sección 3.1.

Así, la distancia basada en el coste extiende a la distancia euclídea y potencialmente permite aplicar los métodos geoestadísticos en situaciones con regiones tanto homogéneas como heterogéneas. En particular, regiones que contengan barreras, como en el proyecto piloto.

3.2.2 Relación con la distancia euclídea

Es natural preguntarse si la utilización de distancias basadas en el coste en lugar de la distancia euclídea tiene un efecto significativo o no. Dependerá claramente de cuán *diferente* (en algún sentido) sea la superficie de coste respecto de una superficie constante, lo que a su vez dependerá de la *geometría* del entorno.

En un entorno urbano, las edificaciones y la infraestructuras urbanas pueden considerarse como barreras arquitectónicas para muchos fenómenos. Contaminantes, ruido, luz, o en general cualquier fenómeno que se propague por el aire y sea bloqueado por los muros son ejemplos claros.

En muchas configuraciones las distancias euclídea y la basada en el coste no serán muy diferentes. Por ejemplo, para cada par de puntos situados a lo largo de una misma vía las dos distancias coincidirán. Sin embargo, no es difícil encontrar situaciones en las cuales los dos tipos de distancias son radicalmente diferentes.

Ejemplo 3.5. En la Figura 3.7 se representan las distancias basadas en el coste desde un conjunto de puntos hasta el punto rojo central. Concretamente, el conjunto está compuesto por todas aquellas localizaciones situadas a una distancia basada en el coste igual o menor que 150 m. Por lo tanto, se puede considerar que el área coloreada es una especie de *círculo basado en el coste* centrado en el punto rojo y de radio 150 m. Consideramos ambos tipos de distancias entre el punto rojo y los cuatro puntos etiquetados como A , B , C y D . Se puede ver a simple vista que A y B están a la misma distancia euclídea del punto rojo; sin embargo, B está dos veces más lejos que A en términos de

distancia basada en el coste. Recíprocamente, D está dos veces más lejos que C , mientras que se encuentran a la misma distancia basada en el coste del punto rojo.

3.2.3 Validez del modelo

Como se ha mencionado en la Sección 3.1, muchos investigadores han trabajado sobre la utilización de distancias no euclídeas en problemas geoestadísticos. Curriero (2006) demostró que muchos de los modelos de covariograma paramétricos tradicionales no son válidos con distancias no euclídeas a través de un sencillo ejemplo que se reproduce a continuación.

Dadas cuatro observaciones en \mathbb{R}^2 dispuestas en las esquinas de un cuadrado de lado unidad, con la métrica dada por la función de distancia *Manhattan*: $d(\mathbf{x}, \mathbf{y}) = |y_1 - x_1| + |y_2 - x_2|$, siendo x_i, y_i , $i = 1, 2$ las coordenadas de \mathbf{x} e \mathbf{y} en algún *sistema de referencia*.

Se considera entonces un covariograma Gaussiano isotrópico (ver Cap. 2), de modo que $C(\mathbf{s}, \mathbf{s}') = C(h = \|\mathbf{s} - \mathbf{s}'\|) = 20e^{-h^2/4}$, $\forall h \geq 0$, cuyos parámetros característicos son *pepita* $\tau^2 = 0$; *alféizar* $\sigma^2 = 20$ y *rango efectivo* $\phi = \sqrt{3 \cdot 4} \cong 3,4641$ (ver Fig. 3.8).

Asumiendo un proceso Gaussiano subyacente, con una función de medias cualquiera (digamos, nula) y con esta estructura de covarianzas, entonces el vector aleatorio (Z_1, Z_2, Z_3, Z_4) se distribuye Normalmente con media nula y matriz de varianzas-covarianzas igual a la que resulta de aplicar la función C a las distancias entre observaciones.

$$\mathbf{D} = \begin{pmatrix} 0 & 1 & 1 & 2 \\ 1 & 0 & 2 & 1 \\ 1 & 2 & 0 & 1 \\ 2 & 1 & 1 & 0 \end{pmatrix} \rightsquigarrow \boldsymbol{\Sigma} = \begin{pmatrix} 20,00 & 15,58 & 15,58 & 7,36 \\ 15,58 & 20,00 & 7,36 & 15,58 \\ 15,58 & 7,36 & 20,00 & 15,58 \\ 7,36 & 15,58 & 15,58 & 20,00 \end{pmatrix}$$

Se obtiene así la matriz de varianzas-covarianzas de las observaciones que, para que sea consistente con la hipótesis del proceso Gaussiano subyacente, ha

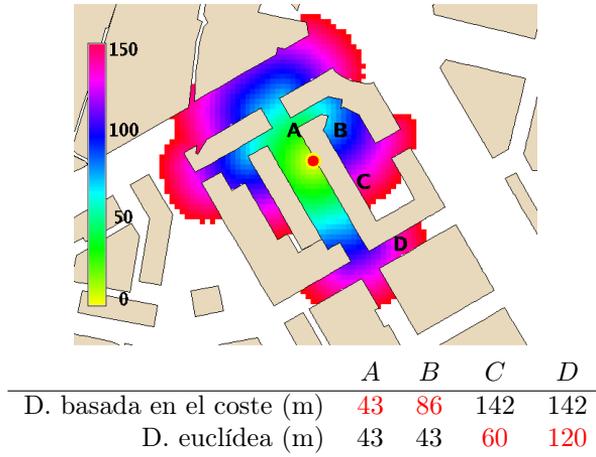


Figura 3.7. Distancias basadas en el coste (hasta 150 m) hasta el punto rojo.

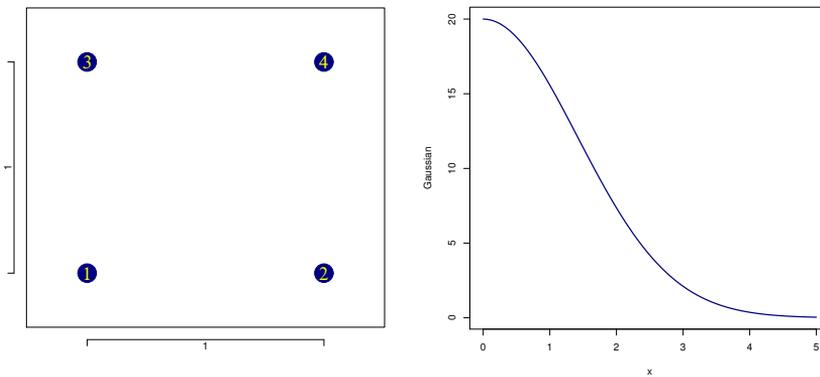


Figura 3.8. Disposición de localizaciones y covariograma en el ejemplo de Curriero (2006).

de ser definida positiva. Sin embargo, los valores propios de esta matriz son $\{58,52; 12,64; 12,64; -3,80\}$.

Curriero (2006) concluye entonces que la función de covariograma Gaussiana no es válida para el uso con la distancia Manhattan. De hecho menciona que tampoco lo son las funciones Esférica, Cuadrático-racional y varias de tipo Matérn. La moraleja es que cuando se usan distancias no euclídeas hay que asegurarse de que la función de covariograma utilizada es definida positiva.

Notar que la distancia Manhattan coincide con la distancia basada en el coste cuando las cuatro localizaciones son las esquinas de una *barrera* cuadrada como puede ser, por ejemplo, un bloque de edificios.

El ejemplo de Curriero es importante para este trabajo porque demuestra, en particular, que cuando se utiliza la distancia basada en el coste no existen garantías de que las funciones de covariograma estándar sean definidas positivas y por tanto el modelo asumido puede ser inválido, al igual que las conclusiones que de él pudieran extraerse.

En la aplicación piloto de este trabajo se ha comprobado que, efectivamente, la matriz de varianzas-covarianzas de las observaciones estimada fuera definida positiva, garantizando así la validez del estudio. Sin embargo, este aspecto abre líneas de trabajo interesantes que nos gustaría explorar en el futuro y que describimos con detalle en el Capítulo 9.

3.3 Cálculo geográfico

3.3.1 Introducción

Las distancias euclídeas se pueden calcular en función de las coordenadas de los puntos. Sin embargo, en la distancia basada en el coste intervienen elementos adicionales, de naturaleza geográfica, como es el caso de la configuración espacial de los edificios en la aplicación piloto. El entorno más natural para el cálculo de la distancia basada en el coste es pues, un *Sistema de Información Geográfica* (GIS, por sus siglas en inglés).

En este capítulo se explica la utilización de herramientas GIS para la predicción geoestadística utilizando distancias basadas en el coste. En particular, el algoritmo de cálculo de las distancias basadas en el coste.

En el Apéndice B se proporcionan los conceptos básicos relacionados con Sistemas de Información Geográfica utilizados en este capítulo. Información más detallada se puede encontrar en la publicación colaborativa y libre de OSGeo ([Open Source Geospatial Foundation, en progreso](#)).

3.3.2 Localizaciones de predicción

El primer producto del proceso es la creación de la capa de localizaciones de predicción, que contendrá los puntos no observados donde se requiere una estimación del fenómeno en cuestión. Esta capa contendrá también la información geográfica asociada a cada localización que será utilizada como covariable en la regresión sobre la media del proceso.

En el caso estándar en el que ningún punto es de particular interés, asumiremos que se requiere predecir sobre todas las áreas donde sea posible, dentro de la región de trabajo; esto es, en todos los sitios donde tenga sentido, como es el caso cuando el objetivo es elaborar un mapa acústico de la zona.

Decidir si predecir en un sitio tiene o no sentido requiere información geográfica, que debe ser proporcionada en la forma de una capa GIS. En el proyecto piloto, la capa con las edificaciones, proporciona información sobre los lugares donde *no* tiene sentido predecir. En otras aplicaciones podría ser más directo obtener una capa con las áreas de predicción, o bien podría ser necesario obtener unas áreas u otras mediante la combinación de varias capas de información.

Cuando se quiere hacer un mapa, en la práctica lo que se hace es representar los valores en una cuadrícula, lo bastante fina como para dar la sensación de continuidad. La resolución de esta cuadrícula es un parámetro del proceso que hay que decidir a priori. Esto determinará la resolución de los mapas de predicción finales.

Con esta resolución se define una cuadrícula de tipo *vectorial* que cubra toda la región. Esto es una función estándar en cualquier GIS. Ahora utilizamos la capa de entrada, con las áreas de predicción o bien con las de no-predicción. En el primer caso, se toma la intersección de la cuadrícula con el área de predicción, mientras que en el segundo caso se resta el área de no-predicción de ella. Así, se obtiene un mosaico de las áreas de predicción, a partir del cual se pueden tomar los centroides de los polígonos como las localizaciones de predicción.

3.3.3 Mapas de distancias

Para el cálculo geoestadístico hacen falta dos matrices de distancias. Una, que contiene las distancias entre las observaciones, será una matriz cuadrada y simétrica con entradas estrictamente positivas, excepto en la diagonal que serán nulas. La segunda matriz contiene las distancias entre las observaciones y las localizaciones de predicción; por lo que se trata de una matriz de tamaño $n(\text{observaciones}) \times m(\text{localizaciones de predicción})$.

El cálculo de ambas matrices se reduce al problema general de calcular la matriz de distancias entre dos conjuntos de puntos cualesquiera, que llamaremos conjunto de *salida* y de *llegada* (sin importar, en principio, cuál es cual). Para el cálculo de la primera matriz, los conjuntos de *salida* y de *llegada* son ambos el conjunto de las observaciones, mientras que para la segunda matriz serán el conjunto de observaciones y de localizaciones de predicción respectivamente.

El cálculo de una matriz de distancias euclídeas es una simple operación aritmética. Pero las distancias basadas en el coste involucran una superficie de coste sobre la que hay que encontrar una ruta óptima para cada pareja de puntos. Se trata de una operación no trivial, que requiere una estrategia computacionalmente eficiente.

Nuestra solución para este problema se basa en el cálculo de un *mapa de distancias* para cada punto en el conjunto de *salida*.

Definición 3.6 (Mapa de distancias asociado a un punto y a una superficie de coste). Se trata de un mapa que representa, para cada localización de la región, la distancia basada en el coste entre ella y el punto dado.

El mapa coloreado en la Figura 3.7 es una representación parcial del mapa de distancias asociado al punto rojo. De esta manera es posible saber la distancia basada en el coste para los puntos *A*, *B*, *C* y *D* simplemente mirando el valor del mapa de distancias en las posiciones respectivas.

Los mapas de distancias son *superficies de coste acumulado*, como se suelen llamar en la jerga de GIS. La mayoría de los Sistemas de Información Geográfica implementan algoritmos para el cálculo de superficies de coste acumulado, que es el paso más costoso en términos computacionales. Tal como concluyen Little et al. (1997, p. 10) luego de programar rutinas en FORTRAN para este paso: «*Integration... with GIS analysis is essential*».

Aprovechando la funcionalidad GIS para el cálculo de superficies de coste acumulado, la operación de cálculo de la matriz de distancias basadas en el coste entre dos conjuntos de puntos se puede automatizar con el siguiente algoritmo.

Algoritmo de cálculo

Problema Calcular la matriz de distancias basadas en el coste entre un conjunto de puntos de *salida* y uno o más conjuntos de puntos de *llegada*

Entradas

- **Un Sistema de Referencia Coordinado (CRS) y una región de trabajo.** Cada entidad del modelo debe estar referenciada en el mismo sistema, para ser capaces de medir distancias relativas. Además, se requiere una región de trabajo *finita*, donde las observaciones y localizaciones de predicción estén contenidas.
- **Las coordenadas de los puntos en los conjuntos de *salida* y de *llegada*.**

- **La superficie de coste.** Esto es, un mapa *ráster* (con una resolución adecuada) cubriendo la región de trabajo.

El funcionamiento es el siguiente: se parte de una capa vectorial de puntos de *salida* y de una o más capas también de puntos, de *llegada*. Además se proporciona un mapa *ráster* *coste*, que representa la superficie de coste. Para cada punto en el conjunto de *salida*, se genera un mapa *ráster* auxiliar de coste acumulado (el mapa de distancias), que cubre toda la región de trabajo. Para cada conjunto de *llegada*, se generan tantas columnas en su tabla de atributos correspondiente, como puntos haya en el conjunto de *salida*. Finalmente, para cada punto de cada conjunto de *llegada*, se toma el valor en la posición correspondiente a dicho punto, de cada uno de los mapas de distancias auxiliares y se escribe dicho valor en la columna correspondiente al punto asociado al mapa de distancias. Como resultado, se obtiene en la tabla de atributos de los mapas de *llegada*, la matriz de distancias de cada punto, a cada uno de los puntos del mapa de *salida*.

En resumen:

PARA cada punto A en *salida*:

- Calcular su *mapa de distancias* D
- PARA cada punto B en *llegada*:

Recoger el valor de D en B

Con la implementación de este sencillo algoritmo en cualquier aplicación GIS y utilizando la capa de observaciones como el conjunto de *salida* y las capas de observaciones (otra vez) y de localizaciones de predicción como conjuntos de *llegada*, se obtienen las dos matrices de distancias basadas en el coste requeridas con un único proceso completamente automático.

3.3.4 Incorporación de información adicional

Los cálculos geoestadísticos pueden realizarse conjuntamente con modelos de regresión sobre la media del proceso subyacente —lo que se conoce como Kriging Universal (ver Cap. 2)—, permitiendo la utilización de variables explicativas que pueden proporcionar información adicional sobre la variable respuesta. Para muchas variables medioambientales, por ejemplo, la altitud resulta ser una covariable muy informativa.

Las distancias a objetos o lugares relevantes pueden proporcionar también información valiosa en algunos casos. Por ejemplo, la distancia a una fuente de vertidos tóxicos en problemas de contaminación, o la distancia a fuentes de ruido en problemas acústicos. Por los motivos desarrollados en la Sección 3.1, en algunas aplicaciones es razonable plantear el modelo en términos de distancias basadas en el coste.

Es necesario algo de cálculo geográfico para hacer esto posible. Debe calcularse un mapa de distancias para cada entidad de interés, para posteriormente, recoger los valores de distancias para cada observación y localización de predicción mediante un proceso iterativo análogo al descrito en la Sección 3.3.3.

3.3.5 Representación de resultados

Una etapa final del proceso donde el cálculo geográfico es especialmente útil es, por supuesto, la representación de los mapas resultantes. El producto de las técnicas geoestadísticas es el conjunto de los valores de predicción y estimaciones de su errores para cada localización de predicción. Estos valores deben ser devueltos al GIS como dos nuevos atributos de la capa de localizaciones de predicción. Puesto que cada localización de predicción es originada como el centroide de un polígono que era parte de un mosaico del área de predicción, tiene sentido asignarle estos resultados a cada uno de los polígonos correspondientes. De este modo se puede generar un mapa continuo de toda la zona de predicción, basado en una escala de color adecuada. Una vez más, ambos procedimientos son funciones típicas de cualquier Sistema de Información Geográfica.

3.4 Conclusiones

En este capítulo hemos dado una respuesta práctica y aplicada a un problema de predicción espacial en una región heterogénea, desarrollando una metodología general para este tipo de escenarios.

El problema específico que hemos utilizado como piloto consiste en la elaboración de mapas acústicos a partir de mediciones realizadas sobre el terreno. Esto puede tratarse como un problema de predicción estadístico en el cual se pretende determinar el valor del nivel de ruido en localizaciones no observadas; concretamente, en una cuadrícula de puntos de predicción cubriendo la región de interés.

Tratándose de un fenómeno espacial continuo, la Geoestadística proporciona las herramientas idóneas para abordar el problema, modelando el fenómeno como una realización de un proceso estocástico espacial, caracterizado por sus funciones de media y de covarianza.

Se han investigado las particularidades del problema en entornos urbanos, demostrando que la aplicación directa de la metodología geoestadística produce resultados poco satisfactorios. La causa fundamental es la presencia de barreras arquitectónicas, que introducen un elemento adicional en la relación entre correlación estadística y distancia.

Se ha diseñado una adaptación de la metodología geoestadística, utilizando el concepto de distancia basada en el coste. Esta solución no es específica del problema de mapas acústicos, sino que es una generalización de la metodología clásica que amplía las posibilidades de aplicación de los métodos geoestadísticos a múltiples situaciones de predicción sobre regiones heterogéneas.

La implementación de la metodología involucra dos operaciones no triviales: el cálculo de las distancias basadas en el coste y la adaptación de los algoritmos geoestadísticos clásicos. Entre este capítulo y el siguiente damos solución a ambas operaciones. Aquí hemos proporcionando un algoritmo general para el cálculo de distancias basadas en el coste cuya implementación detallada se describe en el [Capítulo 4](#).

Aparte de las complejidades de implementación práctica, a las que se ha dado solución entre este capítulo y en el siguiente, la utilización de distancias basadas en el coste reviste una limitación de orden teórico. A saber, cuando se utiliza una distancia no Euclídea no existen garantías de que los modelos de semivariograma clásicos produzcan una estructura de segundo orden válida para el proceso subyacente.

En este punto, queda manifiesta la necesidad de comprobar, en cada aplicación práctica que la matriz de covarianzas estimada para las observaciones sea definida positiva. Si lo es, los resultados son correctos bajo dicho modelo y con ese conjunto de datos particular, aunque el modelo en sí no sea necesariamente válido en general, como se ha discutido en la Sección 2.4.

Sin embargo, como hemos mostrado en la Sección 3.2.3, puede ocurrir lo contrario. Entonces, ¿es posible encontrar condiciones (sobre la superficie de coste, la configuración del entorno y/o el covariograma) bajo las cuales se pueda garantizar la validez del modelo? Mejor aún, ¿es posible encontrar una familia de funciones definidas positivas para la distancia basada en el coste? En ese caso, deberían ser un subconjunto de las existentes puesto que la distancia Euclídea es un caso particular de la distancia basada en el coste. ¿O será que para cualquier función que se proponga es posible encontrar una superficie de coste bajo la cual dicha función no sea definida positiva?

Este problema y estas preguntas constituyen el núcleo de la parte teórica de presente Tesis y en torno a él se desarrollan los Capítulos 5, 6, 7 y 8.

Capítulo 4

Implementación Computacional

En este capítulo se proporcionan los detalles de la implementación específica de la metodología *basada en el coste* introducida en el capítulo anterior. También se presentan y analizan los resultados de la aplicación de dicha metodología al proyecto piloto de elaboración de mapas acústicos en entornos urbanos.

Por su carácter técnico y circunstancial, este capítulo puede ser omitido o pospuesto por el lector interesado en el tema central de esta Tesis.

4.1 Plataforma de trabajo

La utilización de una herramienta GIS combinada con los métodos geostatísticos constituyen el pilar fundamental de la metodología expuesta en este trabajo.

Aunque la metodología es en principio independiente del sistema utilizado, ya que utiliza funciones estándar, la implementación concreta requiere elegir una alternativa en la cual se programen los algoritmos y funciones necesarias.

Hemos utilizado GRASS GIS (GRASS Development Team, 2008) para el análisis geográfico y R (R Development Core Team, 2008) para los cálculos geostatísticos. Ambos son aplicaciones de *Software Libre*, altamente programables y flexibles. Además se comunican fácilmente entre ellos a través de una librería de R llamada `spgrass6` (Bivand, 2008) y permiten ejecutar instrucciones en bloque (*scripts*) lo que facilita la automatización de procesos.

En aplicaciones GIS el acceso libre al código fuente de la aplicación es particularmente importante porque los algoritmos subyacentes pueden ser complejos, y el modo en que han sido implementados puede influenciar enormemente los resultados del modelado espacial y su análisis (Neteler y Mitasova, 2002).

Otros candidatos que hemos considerado viables para la fase de GIS han sido gvSIG (gvSIG Development Team, 2008) y SAGA (SAGA Development Team, 2008).

El paquete `geoR` (Ribeiro y Diggle, 2001) implementa la mayoría de los métodos geostatísticos clásicos en R. Hemos adaptado algunas de sus funciones para implementar los modelos geostatísticos con distancias no euclídeas. Estas adaptaciones se describen con detalle en el Apéndice D, junto con instrucciones para la descarga y utilización del paquete adaptado y compilado.

4.2 Datos iniciales

Como se ha señalado anteriormente, en primer lugar se ha de tener toda la información de interés georeferenciada en un mismo sistema de coordenadas.

Para nuestra aplicación piloto, la información geográfica de interés consiste en un mapa con las manzanas de la zona y una capa con las vías de tráfico. Los edificios se consideran como barreras para el ruido. Aunque las manzanas no necesariamente están edificadas completamente, es la mejor aproximación de la que disponemos, y a los efectos experimentales es igual de útil. Ambos mapas en el mismo sistema de referencia, compuesto por el *Datum* ED50 y *proyección* UTM, huso 30N. La región de trabajo es un rectángulo que contiene a todas las

observaciones dejando un margen amplio, definida exactamente por los rangos de coordenadas entre 4370101,76162 y 4371130,22908 en la dirección Norte–Sur, y 725101,62531 y 726474,309026 en la dirección Este–Oeste.

Las observaciones se han georeferenciado manualmente a partir de las notas tomadas en campo, según su posición relativa con respecto a elementos identificables en los mapas. Aunque en este punto la precisión no es un aspecto crítico, idealmente las mediciones habrían de hacerse con un GPS.

4.3 Localizaciones de predicción

En nuestra aplicación piloto, tenemos una capa con las edificaciones en la zona, que constituyen las áreas donde no tiene sentido predecir. Así que, de acuerdo con el procedimiento descrito en la Sección 3.3.2, comenzamos creando una cuadrícula vectorial con la resolución seleccionada, a la que le restamos la capa de edificaciones. Los centroides de los polígonos resultantes constituyen nuestro conjunto de localizaciones de predicción.

La Figura 4.1 muestra una porción del mosaico del área de predicción con sus centroides, una vez que las edificaciones (en color sólido) han sido sustraídas. Puede observarse que este proceso genera, de forma automática, un conjunto de localizaciones a la resolución deseada o mayor, en aquellos casos situados cerca de los edificios.

Notar que las operaciones llevadas a cabo son funciones típicas de cualquier GIS: creación de una cuadrícula vectorial, intersección/sustracción respecto de otra capa vectorial y extracción de centroides.

4.4 Cálculo de las matrices de distancias

El elemento clave para el cálculo de las distancias basadas en el coste es la superficie de coste. En nuestra aplicación a la elaboración de mapas acústicos, las heterogeneidades de la región de trabajo se reducen a barreras arquitectónicas,

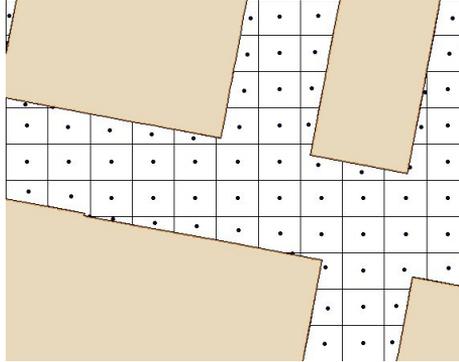


Figura 4.1. Generación automática de las localizaciones de predicción.

que se asumirá que bloquean por completo el paso del ruido. De modo que la superficie de coste debe representar un coste infinito, o muy elevado, en las áreas donde hay edificios y un valor constante en las áreas abiertas por donde el ruido se propaga. Este valor constante representa la distancia, o el coste, de atravesar cada celda del mapa ráster.

En el algoritmo presentado en la Sección 3.3.3 se utiliza una función estándar de los Sistemas de Información Geográfica para el cálculo de superficies de coste acumulado. En GRASS esta función se llama `r.cost` (Awaida y Westervelt, 2006). En `r.cost` los valores nulos son interpretados como coste infinito, por lo que proporciona una forma de representar las barreras.

La resolución del mapa ráster que representa la superficie de coste es un parámetro del proceso de cálculo que afecta tanto a su rapidez como a su precisión, en sentidos opuestos. Teóricamente, la superficie de coste es continua, pero en la práctica es necesario discretizarla, y esto introduce un cierto error. A mayor resolución, menor error. Por otra parte, consume más tiempo calcular los mapas de distancias y éstos ocupan más espacio en el disco duro. Puesto que se requieren tantos mapas de distancias como observaciones, la diferencia en tiempo de proceso y en espacio en el disco duro puede ser muy significativa. Así,

la resolución es un parámetro de compromiso, cuyo valor depende en particular de n , el número de observaciones.

El valor constante no nulo asignado a la celda ráster en la superficie de coste discretizada representa la integración de la superficie continua ideal a lo largo de una curva que atraviesa esa celda específica. Por tanto, el valor asignado sólo puede ser una aproximación a la distancia recorrida al atravesar la celda y por tanto depende de su tamaño; es decir, de la resolución del mapa ráster. Obviamente, el valor también depende de la dirección relativa en la cual se atraviese la celda. Sin embargo, este factor se puede corregir en el proceso de cálculo del mapa de distancias, como se ve a continuación.

En nuestra aplicación la resolución utilizada ha sido de 5 m, así que cada celda representa una porción de terreno de $5\text{ m} \times 5\text{ m}$, y el coste asignado a aquellas que no constituyen barreras también es de 5; lo que significa que atravesar la celda tiene un coste equivalente a recorrer 5 m.

Los mapas de distancias están implementados a través de superficies de coste acumulado y se calculan por *expansión* desde el punto al que están asociados y por acumulación de los costes de cada celda. Existen varias estrategias de expansión posibles. Por ejemplo, en cada paso la expansión puede producirse a las celdas adyacentes, o también en diagonal, multiplicando el coste de la celda por un factor de corrección. Aún más sofisticado es expandir además al estilo del Caballo del Ajedrez, lo que incrementa la precisión ostensiblemente. Sobre una superficie de coste constante, el coste acumulado se expande en 4, 8 o 16 direcciones respectivamente (ver Fig. 4.2).

Hemos desarrollado un script de GRASS que implementa de forma automática este procedimiento de cálculo de las matrices de distancias. En el Apéndice C se proporcionan sus instrucciones de uso, su código fuente, e instrucciones para su descarga.

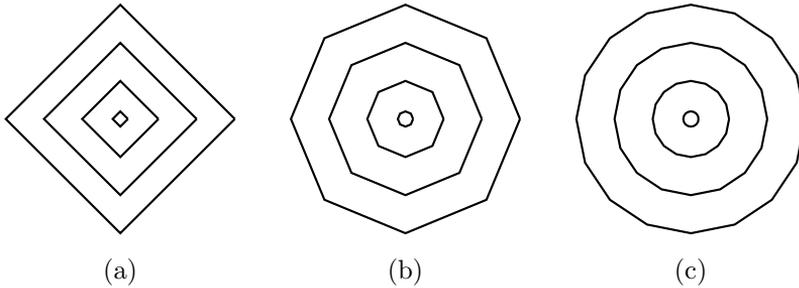


Figura 4.2. Curvas de nivel de coste acumulado para una superficie de coste constante con distintas estrategias de expansión. (a) Expansión hacia celdas adyacentes únicamente; (b) expansión adicional en diagonal; (c) expansión adicional según el movimiento del Caballo del Ajedrez.

4.5 Información adicional

El ruido generado por el tráfico rodado es habitualmente considerado como la principal fuente de ruido urbano. Por tanto las diferencias en la densidad del tráfico explican gran parte de las diferencias en los niveles de ruido.

Por este motivo se han clasificado las vías de acuerdo con la densidad de tráfico media estimada. Las avenidas se consideran vías de categoría 1, mientras que a las calles con densidad de tráfico media y baja se les ha asignado las categorías 2 y 3 respectivamente.

En una primera aproximación, esta información se puede incorporar directamente como factor, contribuyendo en distinta medida al nivel de ruido de una localización según la categoría de la vía sobre la que esté ubicada, tal como se ha hecho en la Sección 3.1. Sin embargo ésta es una aproximación demasiado grosera. Las intersecciones, o las zonas alejadas de cualquier vía son áreas que presentarían discontinuidades espaciales difícilmente justificables. Por otra parte, puntos situados en una vía de baja intensidad de tráfico tendrían la misma contribución independientemente de lo cerca o lejos que estén de una vía de alta intensidad de tráfico.

En cambio, parece mucho más informativo considerar la distancia hasta la avenida más cercana. Así, aquellos puntos situados sobre la avenida propiamente dicha tendrían una distancia cero, lo que implicaría una determinada contribución a la variable respuesta que iría disminuyendo suavemente conforme el punto se aleja de la avenida. Por supuesto tiene sentido considerar —aquí también— distancias basadas en el coste en lugar de distancias euclídeas.

De modo que, para cada punto —observaciones y localizaciones de predicción— se ha de calcular las distancias basadas en el coste hasta las vías de tipo 1, 2 y 3 respectivamente más cercanas. A estos efectos se aplica un procedimiento iterativo análogo al descrito en la Sección 3.3.3.

Hacemos hincapié en el uso de la distancia basada en el coste en lugar de la distancia euclídea para la incorporación de variables explicativas, ya que hay configuraciones en las cuales un punto está muy cerca de la avenida, pero está protegido por una edificación que evita que el ruido le alcance.

La capa de vías de tráfico se ha procesado asignando a cada vía su categoría y separándola en tres capas distintas, cada una conteniendo las vías de una categoría. En primer lugar, puesto que las vías están representadas como líneas en las capas vectoriales, se ha generado un área de influencia (*buffer*) aproximada alrededor de la línea, que representará el ancho de la vía. Se ha utilizado un radio de buffer de 10, 8 y 6 m para las vías de tipos 1, 2 y 3, respectivamente. Luego se han rasterizado estas áreas y se han utilizado como origen para la función de expansión `r.cost`. Así se obtienen los mapas de distancias hasta cada tipo de vía, habiéndose calculado hasta una distancia máxima de 100 m, más allá de la cual se considera que el ruido producido por la vía ya no tiene un efecto significativo (ver Fig. 4.3).

Finalmente, sólo queda recorrer las observaciones y las localizaciones de predicción y recoger los valores de las distancias basadas en el coste hasta cada tipo de vía en las correspondientes columnas de sus respectivas tablas de atributos.

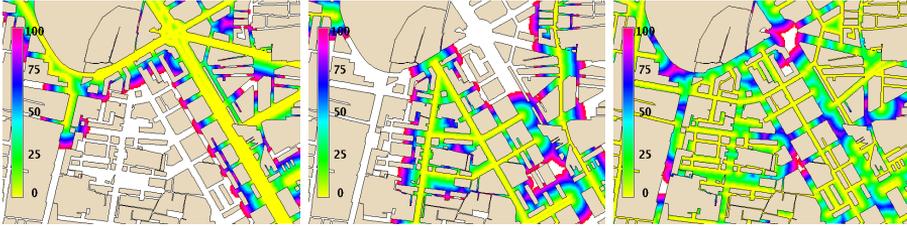


Figura 4.3. Mapas parciales de distancias basadas en el coste hasta la vía más cercana, de cada categoría.

4.6 Cálculos geoestadísticos

Una vez calculadas las matrices de distancias basadas en el coste y las distancias basadas en el coste que intervienen como covariables en el modelo, es necesario transferir esta información al entorno de cálculo estadístico. Desde dentro de la consola de comandos de GRASS se puede abrir una sesión de R y ejecutar scripts, que de este modo tienen acceso a las variables de entorno definidas en GRASS. En particular, a través de la librería `spgrass6` se pueden transferir las capas de observaciones y de localizaciones de predicción a los correspondientes objetos espaciales de R, junto con sus tablas de atributos conteniendo toda la información sobre las distancias basadas en el coste.

El procesamiento estadístico se divide en cuatro etapas: la depuración de los datos; la construcción del modelo de regresión sobre la media del proceso; el ajuste del modelo de semivariograma; y, finalmente, la predicción Kriging propiamente dicha. En esta sección describimos estas etapas con detalle.

4.6.1 Depuración de los datos

En primera instancia, es necesario realizar una depuración previa de los datos, para corregir errores de implementación propios del proceso o de limitaciones de los programas.

Los valores nulos en las tablas de atributos de capas vectoriales de GRASS se pasan como ceros a los objetos resultantes en R, lo que los hace indistinguibles

de los ceros originales. Esto se soluciona distinguiéndolos de alguna manera justo antes de la transferencia y deshaciendo el cambio justo después.

Como consecuencia de la discretización del área de predicción o simplemente por la configuración de la zona, es posible que queden localizaciones de predicción aisladas; es decir, rodeadas de barreras. Esto hace que durante el cálculo de los mapas de distancias, el proceso de expansión nunca las alcance y por tanto estas localizaciones tendrán una distancia basada en el coste desconocida a cualquiera de los demás puntos, u objetos relevantes que se hayan utilizado como covariables. Utilizar las distancias euclídeas en estos casos es una mala opción, ya que precisamente por estar aislados, resultan ser del tipo de caso patológico en los cuales la predicción geoestadística estándar proporciona predicciones absurdas. Por lo tanto, es mejor excluirlos del estudio.

Otra consecuencia de la discretización es que podrían quedar dos observaciones en la misma celda de la superficie de coste, en caso de que estuvieran suficientemente cerca (ver Fig. 4.4a). Esto hace que la distancia basada en el coste ϑ entre las dos observaciones sea cero y que la matriz de varianzas covarianzas del modelo subyacente sea singular, generando un error en el algoritmo de predicción. Con el mismo criterio, sabemos que la distancia basada en el coste debe ser necesariamente mayor que la euclídea, lo cual puede no verificarse en distancias pequeñas como consecuencia del error introducido en la discretización. En estos casos es conveniente sustituir la distancia basada en el coste calculada, por la distancia euclídea correspondiente.

Cuando las distancias son comparables con la resolución utilizada en la discretización, también puede producirse el efecto contrario: que dos localizaciones queden dispuestas de tal manera que la distancia basada en el coste sea mucho mayor que la euclídea, siendo que no hay ningún obstáculo de por medio (ver Fig. 4.4b). Se han controlado las parejas de observaciones situadas a menos de 10 veces la resolución de discretización y que están en el decil superior respecto al cociente entre la distancia basada en el coste y la euclídea. Se han encontrado así dos parejas que efectivamente no tenían obstáculos entre ellas y cuya distancia

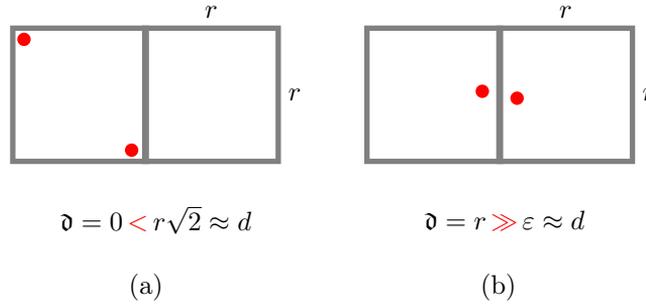


Figura 4.4. La discretización del área de predicción con una resolución r induce artificialmente, entre los dos tipos de distancias, diferencias significativas que es necesario corregir.

basada en el coste se ha corregido a la baja. Lo mismo se ha hecho con las parejas de observaciones y localizaciones de predicción.

4.6.2 Modelo de regresión sobre la media del proceso

Una vez depurados los datos es necesario especificar el modelo de regresión no espacial, para las covariables disponibles. Ésta es una tarea difícilmente automatizable, ya que depende del tipo de covariables específicas que se utilice. Hay que analizar si requieren una transformación previa, las posibles interacciones, multicolinealidad, selección de variables, etc. Es el trabajo típico que requiere cualquier modelo de regresión. Sin embargo, una vez definido, basta introducirlo en el proceso sin más intervención por parte del usuario.

4.6.3 Modelo de semivariograma

Análogamente, el modelo de semivariograma constituye la componente espacial del modelo y también requiere de la intervención de un especialista. Estas componentes del modelo deben analizarse y definirse una vez calculadas las distancias basadas en el coste. Por lo que es necesario hacer una interrupción en el procedimiento y continuar con el proceso una vez definido completamente el

modelo. Es posible que el modelo sirva para otro banco de datos con características similares, en cuyo caso el proceso podría lanzarse de forma automática sin interrupciones.

En el análisis geoestadístico clásico hay tres grandes etapas que deben ser adaptadas para utilizar distancias basadas en el coste: el cálculo del semivariograma empírico, el ajuste del modelo de semivariograma paramétrico y la propia predicción Kriging. Proporcionamos a continuación un resumen conceptual de estas adaptaciones, que se describen con detalle en el Apéndice D.

El semivariograma empírico se calcula únicamente a partir de los datos observados. Es la herramienta que permite hacer una inspección visual de los datos y un diagnóstico de los posibles modelos de semivariograma ajustados. Se clasifican las parejas de observaciones en grupos, según la distancia que las separa, y luego se calcula un estimador del valor teórico del semivariograma, para esa distancia, basado en las diferencias entre los valores observados (ver Sec. 2.2). A los efectos de calcular un semivariograma empírico basado en el coste basta con hacer la clasificación inicial en función de las distancias basadas en el coste dadas en la correspondiente matriz, en lugar de hacerlo en función de las distancias euclídeas. Notar que esta modificación produce un agrupamiento de observaciones diferente y por tanto las estimaciones del valor teórico del semivariograma serán también diferentes. En `geoR` el semivariograma empírico se calcula con la función `variog()`, a la que se le ha añadido un argumento opcional extra que permite pasar una matriz de distancias proporcionada por el usuario. Se ha agregado el código de control necesario para dejar que actúe como la función original en caso de omisión del argumento, pero que realice la clasificación a partir de los datos contenidos en la matriz en caso contrario.

El ajuste del modelo de semivariograma paramétrico también se hace utilizando únicamente los datos observados. Típicamente se utilizan métodos de máxima verosimilitud, a través de algoritmos iterativos. Esto implica el cálculo de la matriz de covarianzas para cada combinación de parámetros ensayada. Todo lo que necesitamos es asegurarnos de que la matriz de covarianzas se calcule a partir de las distancias basadas en el coste. En `geoR` el modelo de semivariograma

se ajusta conjuntamente con la componente no espacial a través de la función `likfit()`, a la cual se le ha añadido un argumento opcional extra análogo al de la función `variog()`, para proporcionar una matriz de distancias no euclídeas.

4.6.4 Predicción Kriging

En este punto el modelo de covarianzas se asume conocido. Sin embargo, aquí también es necesario asegurarse de que la matriz de covarianzas entre las observaciones se calcule en función de las distancias basadas en el coste proporcionadas por el usuario. Además, las covarianzas entre las observaciones y las localizaciones de predicción deben ser calculadas, y aquí es donde es necesario utilizar la segunda matriz de distancias basadas en el coste. En `geoR` la predicción Kriging se realiza con la función `krige.conv()`, a la cual se le han añadido los dos argumentos opcionales necesarios para pasar las dos matrices de distancias basadas en el coste, además del resto de modificaciones pertinentes para asegurar el funcionamiento normal en caso de omisión y el funcionamiento adecuado cuando se proporcionen las matrices por el usuario (ver Ap. D).

4.7 Esquema global del proceso

El proceso completo se esboza a continuación para el caso particular de regiones con barreras, como es el caso en la aplicación a los mapas acústicos.

Cálculo geográfico

Entradas: entorno geográfico, mapa de barreras, parámetro de resolución de predicción.

1. Crear una cuadrícula vectorial cubriendo la región de trabajo con la resolución especificada.
2. Recortar las áreas donde existen barreras. El resultado es un mosaico del área de predicción.

3. Extraer como localizaciones de predicción los centroides de cada polígono del mosaico.
4. Incorporar información adicional como covariables: distancias basadas en el coste hasta entidades relevantes. Iterar sobre cada una de las entidades de interés. *Entradas: mapas de entidades y de área de predicción, mapas de observaciones y de localizaciones de predicción, parámetros de resolución para el cálculo del coste, y coste máximo considerado.*
 - 4.1. Rasterización del mapa de la entidad con la resolución dada.
 - 4.2. Rasterización de la región de predicción con la resolución dada y con un valor igual al tamaño de la celda. El resultado es la superficie de coste.
 - 4.3. Cálculo del mapa de distancias desde la entidad hasta el coste máximo.
 - 4.4. Recoger como atributos las distancias basadas en el coste para las observaciones y localizaciones de predicción.
5. Calcular las matrices de distancias basadas en el coste: observaciones-observaciones y observaciones-localizaciones de predicción. *Entradas: mapas de observaciones y localizaciones de predicción, superficie de coste (mapa de barreras).*
 - 5.1. Calcular los mapas de distancias para cada una de las observaciones.
 - 5.2. De cada mapa de distancias, recoger como atributos las distancias basadas en el coste para las observaciones y localizaciones de predicción.

Predicción geoestadística

Entradas: mapas de observaciones y de localizaciones de predicción, con sus tablas de atributos conteniendo valores de covariables y matrices de distancias basadas en el coste.

1. Selección del modelo de regresión. Transformación y selección de covariables, interacciones, etc.
2. Cálculo del semivariograma empírico basado en el coste.
3. Selección y estimación del modelo de semivariograma.
4. Predicción Kriging, basada en el coste.
5. Devolver el mapa de localizaciones de predicción al GIS con las predicciones y las estimaciones del error añadidas como atributos.

Representación de resultados

Entradas: mapa de localizaciones de predicción conteniendo en sus tablas de atributos los valores predichos y sus errores estimados, mosaico de la región de predicción, elementos adicionales que se deseen representar.

1. Transferir los valores predichos y sus errores estimados desde las localizaciones hasta los polígonos correspondientes en el mosaico de la región de predicción.
2. Configurar las opciones de los mapas temáticos y mostrar resultados.

4.8 Aplicación al proyecto piloto

4.8.1 Comparación de ambos tipos de distancias

A pesar de que en muchas situaciones las distancias basada en el coste y euclídea serán similares, o incluso iguales, como cuando los puntos están situados a lo largo de una misma vía, en la Sección 3.2.2 mostramos que en algunos casos los dos tipos de distancias pueden ser radicalmente diferentes.

En la Figura 4.5 se comparan los valores de los dos tipos de distancias para todas las parejas de observaciones en la aplicación piloto.

Hay un umbral de distancia (Euclídea), alrededor de los 95,5 m, a partir del cual se produce un aumento repentino en la variabilidad de la relación entre los

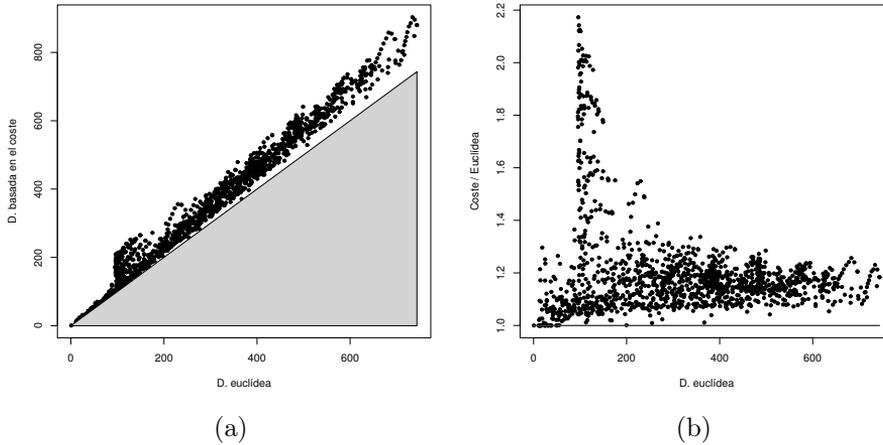


Figura 4.5. Relación entre los dos tipos de distancias para todas las parejas de observaciones.

dos tipos de distancias. Las únicas observaciones que están a menos de 95 m entre sí son las correspondientes a las mediciones más densas, que están sobre la misma vía (ver Fig. 3.4). De modo que las distancias son prácticamente iguales, con una pequeña variabilidad debida a la discretización del espacio introducida para el cálculo aproximado de las distancias de coste.

A partir de ese umbral, ya se encuentran puntos en diversas configuraciones espaciales y la variabilidad es entonces mayor. Se trata, por tanto, de un artefacto producto de la configuración del muestreo. Entre ese punto de salto y los 300 m aproximadamente, hay una gran variabilidad que decae rápidamente de modo que a distancias mayores la relación entre los distintos tipos de distancias se mantiene entre 1,1 y 1,3.

Un enfoque alternativo a la distancia basada en el coste que hemos considerado es la utilización del escalado multidimensional (MDS, ver Sec. 6.1) con el objetivo de obtener una aproximación euclidiana a la estructura de distancias, aunque fuera necesario un espacio de dimensión mayor. Hemos encontrado que con 2

dimensiones conseguimos explicar aproximadamente el 84% de la distribución espacial.

Es interesante contrastar el comportamiento de los tres tipos de distancias. En la Figura 4.6 se puede ver, para cada pareja de puntos en este juego de observaciones, la distancia euclídea que las separa y la aproximación MDS a esta distancia. Las parejas están ordenadas en orden creciente de distancia basada en el coste. Se pueden observar, desde otro punto de vista, los mismos efectos que en la Figura 4.5. Por ejemplo, el incremento notable de la variabilidad a partir de los 95 m, que también se observa en la aproximación MDS. Aunque esta aproximación mantiene la variabilidad con respecto al valor de la distancia basada en el coste parece mantenerse *centrada* alrededor de este valor, corrigiendo la desviación a mediana y gran escala que se produce entre la distancia euclídea y la basada en el coste. En este sentido, podemos concluir que la aproximación MDS corrige la desviación media, a largo plazo, pero no la variabilidad a pequeña escala que es la que produce las incongruencias que han justificado el desarrollo de la distancia basada en el coste.

Además de cuantificar la discrepancia que puede haber entre las distancias euclídea y basada en el coste, es interesante determinar dónde se producen, con respecto a los elementos del entorno. La Figura 4.7 muestra la configuración espacial del 10% de las parejas con mayor discrepancia entre los dos tipos de distancias.

Puede verse que se trata de aquellas observaciones situadas en las regiones más irregulares, o que han sido tomadas lejos de las esquinas. Por el contrario, en el sector Sur-Este, donde adoptan una distribución más regular y las mediciones han sido tomadas cerca de las esquinas, no se observan casos de gran discrepancia.

4.8.2 Modelo de regresión sobre la media

En nuestra aplicación tenemos tres covariables externas, a saber, las distancias basadas en el coste desde cada observación y localización de predicción hasta la vía más cercana de cada una de las tres categorías. Existen motivos físicos

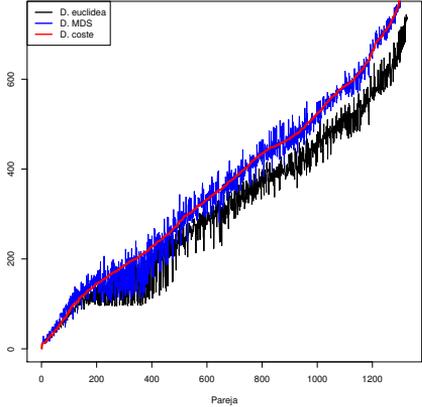


Figura 4.6. Comparación de tipos de distancias para todas las parejas de observaciones.

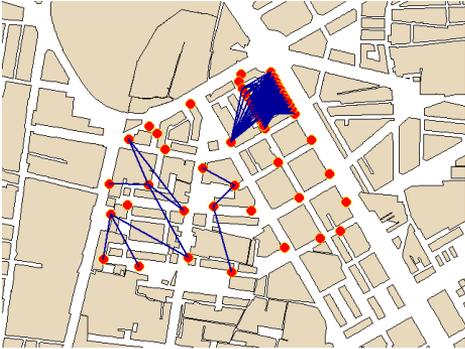


Figura 4.7. Parejas de observaciones con mayores discrepancias entre ambos tipos de distancias.

que justifican una relación lineal entre el nivel de ruido y el inverso del cuadrado de la distancia a la fuente. También resulta bastante evidente que es necesario considerar posibles interacciones entre ellas, ya que por ejemplo, un punto situado en una intersección estará muy cerca de dos vías, pero no por eso el efecto de ambas es aditivo.

Mediante un análisis preliminar se consideró adecuado utilizar una transformación de ellas dada por la función $1/(\frac{x}{20})^2$, e incorporar los efectos principales de las tres covariables transformadas y las interacciones entre la transformada de la distancia hasta la vía de mayor intensidad de tráfico con las otras dos covariables transformadas.

No se ha buscado un modelo de regresión lineal óptimo; por una parte porque esto no garantiza ningún tipo de optimalidad en la estimación de las características del proceso; y por otra parte porque el objetivo central del trabajo se encuentra en la utilización de la distancia basada en el coste, independientemente del modelo de media o covarianza utilizado.

Se realizaron diagnósticos elementales sobre la calidad del modelo (detección de multicolinealidad, significación de los coeficientes, normalidad) y se seleccionó el mejor modelo de entre los ensayados atendiendo a criterios estándar, incluyendo la varianza residual.

4.8.3 Análisis del semivariograma

Las diferencias entre los dos tipos de distancias influye directamente en la estimación de la estructura de correlaciones. En el cálculo del semivariograma empírico, los pares de observaciones separadas por distancias euclídeas similares son agrupadas en el mismo intervalo de distancias. Sin embargo, es muy probable que estén a muy diferentes distancias basadas en el coste y que por tanto contribuyan a explicar partes diferentes de un semivariograma empírico basado en el coste.

La Figura 4.8a muestra los semivariogramas empíricos para ambos tipos de distancias.

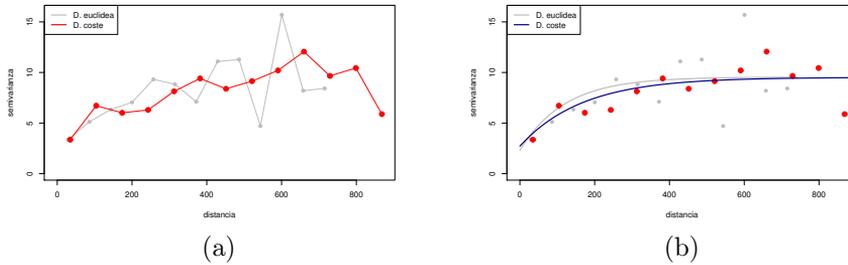


Figura 4.8. Comparación entre los semivariogramas euclídeo y basado en el coste. (a) Semivariogramas empíricos; (b) modelos exponenciales ajustados.

Cuando se utiliza la distancia euclídea, el semivariograma empírico es ostensiblemente más *inestable* o *ruidoso*. Posiblemente porque, al despreciar los obstáculos, está agrupando en el mismo intervalo de distancias parejas de puntos con configuraciones cualitativamente muy diferentes y por tanto con grados de correlación muy distintos, tal como ocurría en el esquema de la Figura 3.6. El hecho de que el semivariograma empírico basado en el coste sea más suave confirma la idea de que las distancias basadas en el coste explican la correlación mucho mejor que las euclídeas.

En esta aplicación, la familia exponencial de semivariogramas fue utilizada para estimar el modelo de semivariograma. Se ha comprobado la validez del modelo resultante, controlando que la matriz de varianzas-covarianzas de las observaciones es efectivamente definida positiva. El Cuadro 4.1 muestra los parámetros ajustados por máxima verosimilitud para ambos tipos de distancias, mientras que los modelos se muestran en la Figura 4.8b. El semivariograma basado en el coste tiene un rango efectivo mayor. En parte, esta diferencia es atribuible al hecho de que las distancias basadas en coste son siempre mayores que las euclídeas.

		Euclídea	Bas. coste
Pepita	$\hat{\tau}^2$	2,30	2,71
Alféizar	$\hat{\sigma}^2$	9,55	9,51
Rango efectivo	$\hat{\phi}$	358	502

Cuadro 4.1. Parámetros ajustados de la familia de semivariogramas exponencial para cada tipo de distancias.

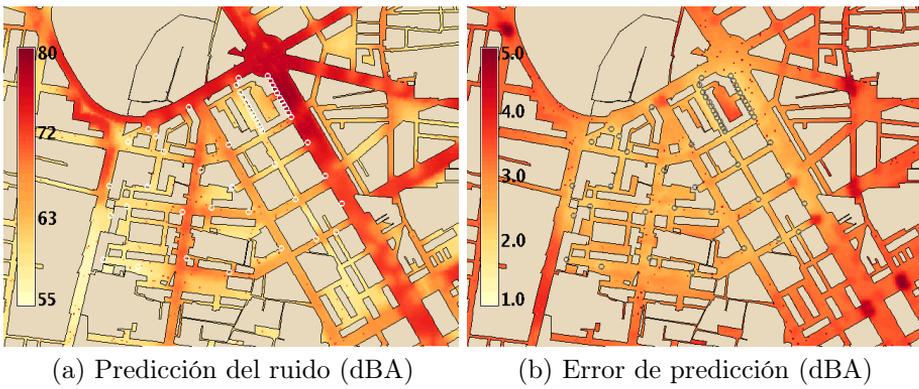


Figura 4.9. Predicción Kriging basada en el coste. Los puntos de medición están señalados ligeramente como referencia.

4.8.4 Resultados de predicción

La Figura 4.9 muestra la salida final del proceso completo: el mapa con las predicciones para cada localización y el mapa con las estimaciones del error estándar de dicha predicción, que es una medida de la incertidumbre de la misma.

Es interesante comparar estos resultados con los que se hubieran obtenido de haber utilizado distancias euclídeas. El mapa tiene prácticamente el mismo aspecto, por lo que es mejor concentrarse en las *diferencias* en las predicciones y en sus errores.

Sin embargo, se debe tener en cuenta que ambos resultados están contruidos con el mismo modelo de regresión sobre la media, con las distancias basadas en el coste hasta las vías más cercanas de cada tipo como covariables. Esto significa que los resultados euclídeos de esta sección no son completamente euclídeos. Pero esto nos permite apreciar las diferencias que el cambio de métrica produce en la estructura espacial exclusivamente.

Las diferencias de predicción entre el Kriging basado en el coste y el Euclídeo para este proyecto piloto van desde $-1,655$ dBA hasta $1,876$ dBA en términos absolutos, y desde $-2,7\%$ hasta $3,1\%$ en términos relativos. En promedio, las diferencias están muy cerca de cero, y en el 95% de las localizaciones son menores que $\pm 1\%$.

Con respecto a la incertidumbre, las diferencias en el error estándar van desde $-0,25$ dBA hasta $0,65$ dBA. En términos relativos, estas diferencias abarcan el rango mucho más significativo de $-5,8\%$ a $27,1\%$. En el 77% de las localizaciones, la predicción basada en el coste es más precisa que la euclídea, pero hay algunas localizaciones donde la incertidumbre es mucho mayor, alcanzando hasta un 27% de error adicional (ver Fig. 4.10).

Lo que resulta más interesante es la distribución espacial de estas diferencias, a los efectos de interpretar en qué situaciones las dos estrategias divergen (ver Fig. 4.11).

Notar que la mayor diferencia se produce en la región interior de la manzana en la esquina superior, donde la estrategia basada en el coste predice un nivel de

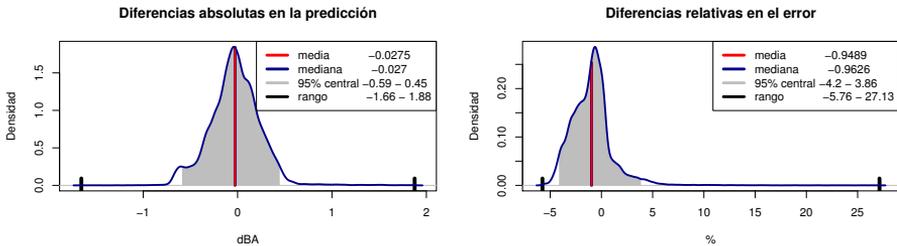


Figura 4.10. Resúmenes muestrales y estimaciones kernel de las distribuciones de las diferencias entre las predicciones basadas en el coste y las euclídeas, y sus errores.

ruido mayor, con la mayor diferencia con respecto a la predicción euclídea. Para la estrategia Euclídea, el área *interior* no es tal; por tanto las observaciones en la vía adyacente con baja intensidad de tráfico tienen una mayor influencia de la que deberían en virtud de la presencia del edificio. En cambio, la estrategia basada en el coste *entiende* que el nivel de ruido en esa área depende mucho más del tráfico de la avenida y por tanto predice un nivel de ruido ostensiblemente mayor.

El mapa con las diferencias relativas en los errores de predicción está mayormente en tonos de azul, lo que significa que la estrategia basada en el coste es generalmente más precisa. Una excepción es precisamente el área comentada en el párrafo anterior; lo cual es fácil de explicar, ya que el método euclídeo *piensa* que hay muchas observaciones muy cerca alrededor de esa área, por lo tanto asigna una gran precisión a la predicción en ese sitio, subestimando la incertidumbre. En cambio, el método basado en el coste tiene en cuenta los edificios. Como consecuencia, las observaciones ya no están tan cercanas y por este motivo la incertidumbre es mayor.

Estos resultados deben considerarse con cautela. El objetivo de este trabajo no ha sido elaborar un mapa acústico de la zona. En ese caso, las mediciones hubieran debido hacerse conforme al protocolo establecido, se debía haber comprobado las hipótesis de normalidad, estacionariedad e isotropía, y se debía haber hecho

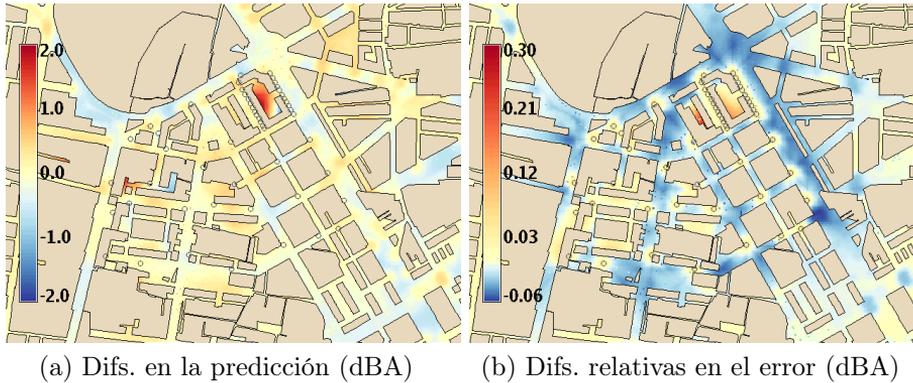


Figura 4.11. Diferencias entre los resultados basados en el coste y los euclídeos. Los puntos de medición están señalados ligeramente como referencia.

un diagnóstico y una validación del modelo. Esto es, la aplicación rigurosa de los procedimientos propios de un trabajo profesional. En cambio, el objetivo ha sido completar una aplicación *piloto*, en la cual nos hemos asegurado de tener unas mediciones y un modelo de una calidad razonable y suficiente como para poder analizar los efectos de las innovaciones introducidas.

4.9 Conclusiones

La implementación de la metodología desarrollada en el capítulo anterior involucra dos operaciones no triviales: el cálculo de las distancias basadas en el coste y la adaptación de los algoritmos geoestadísticos clásicos. En este capítulo proporcionamos los detalles de la implementación específica que hemos utilizado para el proyecto piloto.

En primer lugar justificamos la elección de GRASS y R/geoR como herramientas de desarrollo.

En GRASS se implementó el cálculo de las matrices de distancia basada en el coste, con una función que permite además la incorporación de factores

potencialmente determinantes, concretamente distancias (basadas en el coste) hasta objetos o lugares de especial interés.

La adaptación de los algoritmos geoestadísticos con el objetivo de que admitan la utilización de distancias basadas en el coste se ha realizado sobre el paquete `geoR`. Concretamente se han modificado la función para la construcción de variogramas empíricos; la función para el ajuste máximo-verosímil (conjunto) del modelo de variograma y del modelo lineal de la media; y por último, la función para la predicción Kriging.

Finalmente, presentamos los resultados de la aplicación de esta metodología a los datos del proyecto piloto, comparándolos especialmente con los que se obtienen bajo la utilización de distancias Euclídeas. Hemos encontrado, por término medio, pocas diferencias en las predicciones puntuales, aunque como es natural, muy concentradas en las regiones más irregulares. En cambio las diferencias en el error de predicción llegan a ser bastante importantes. En términos relativos, utilizando distancias Euclídeas se llega a alcanzar hasta un 27 % de error adicional sobre la metodología basada en el coste.

En los Apéndices [C](#) y [D](#) proporcionamos el código fuente de todos los algoritmos y modificaciones implementadas, con enlaces a la web de referencia del proyecto desde donde se pueden descargar, junto con los datos utilizados y código de ejemplo de uso. Con el código de ejemplo publicado se pueden reproducir los resultados de la aplicación al proyecto piloto presentados en este capítulo.

En suma, hemos desarrollado una herramienta práctica que explota la potencia de cálculo geográfico de un GIS y permite obtener predicciones más exactas, proporcionando además, una valoración de la incertidumbre de la misma.

Adicionalmente, la herramienta permite introducir en el modelo factores potencialmente determinantes, concretamente distancias (basadas en el coste) hasta objetos o lugares de especial interés. Esto aporta información valiosa que no podía ser aprovechada antes, permitiendo conocer su aportación a los niveles globales de la variable respuesta, mejorando la precisión del modelo y facilitando la toma de decisiones sobre acciones correctoras.

Por otra parte, en el problema particular de los mapas acústicos en entornos urbanos, creemos que sería muy ventajoso utilizar los resultados de los métodos de cálculo como predictores de la media del proceso.

Los métodos de simulación pueden verse como un caso particular de lo que en la literatura especializada se llama *Computer Models*. Esto es, un programa informático que dados unos parámetros de entrada produce un resultado de forma determinista. Esta es un área de investigación muy activa en los últimos años, gracias al aumento de las capacidades computacionales. Algunas de las líneas de trabajo más importantes en esta área incluyen cómo cuantificar la incertidumbre en parámetros de entrada no observados; cómo trasladar la incertidumbre en los parámetros a los resultados de predicción; cómo tener en cuenta la discrepancia del modelo con la realidad; y especialmente, cómo combinar la predicción del modelo con otras fuentes de información (*Data Assimilation*).

Vale la pena destacar que, dada la complejidad de estos modelos, únicamente bajo el paradigma Bayesiano es posible tener en cuenta todas las incertidumbres adecuadamente. [Kennedy y O'Hagan \(2001\)](#) elaboran un marco metodológico general en el cual trabajar con este tipo de problemas y discuten en particular el tratamiento de la discrepancia del modelo. [Bayarri et al. \(2007\)](#) desarrollan sobre ese marco el tratamiento particular de modelos con salida funcional, y [Bhat et al. \(2010\)](#) discuten en particular el caso espacial, como es el de los simuladores de difusión de ruido.

Esto completaría un sistema integrado de predicción que combinaría de forma armoniosa las distintas herramientas de análisis y predicción disponibles, aprovechando lo mejor de cada una y propiciando una sinergia sin lugar a dudas beneficiosa desde el punto de vista de la predicción. Lamentablemente, la falta de disponibilidad de herramientas de simulación libres (esto es, con una licencia que permita la utilización y modificación del código fuente) entorpece la investigación en este sentido.

En esta aplicación se ha debido comprobar de forma específica que la matriz de covarianzas de los datos resultara definida positiva, con la función de covarianza utilizada. Esto garantiza que los resultados de predicción tienen sentido, en tanto

que el predictor Kriging es el mejor predictor lineal insesgado. Sin embargo la necesidad de realizar esta comprobación en cada caso es muy inconveniente, ya que no hay garantías de validez ni siquiera cuando se incorpore una observación adicional al banco de datos.

En lo que resta de esta Tesis abordaremos el tema central de la permisibilidad de las funciones de covarianza cuando se utilizan distancias no Euclídeas en general y basadas en el coste en particular.

Capítulo 5

Geoestadística Riemanniana

La solución práctica adoptada en los dos capítulos anteriores carece de sustento teórico que la legitime metodológicamente en general. Las funciones de correlación que hemos propuesto utilizar son probadamente válidas en un espacio Euclídeo. En caso de utilizar distancias basadas en el coste no existe ninguna garantía de que la función en cuestión sea definida positiva, aunque circunstancialmente y para un banco de datos determinado sea posible comprobar que el predictor Kriging es válido.

La condición de definición positiva para la función de correlación de un modelo espacial está estrechamente vinculada a la condición de positividad para la varianza de una variable aleatoria. En efecto, dado un campo aleatorio $Z(\mathbf{s})$, para todo n finito, todo conjunto de puntos $\mathbf{s}_1, \dots, \mathbf{s}_n \in D$ y todo conjunto de coeficientes reales a_1, \dots, a_n la combinación lineal $\sum_{i=1}^n a_i Z(\mathbf{s}_i)$ define una variable aleatoria, con una cierta varianza no negativa

$$\mathbb{V} \left[\sum_{i=1}^n a_i Z(\mathbf{s}_i) \right] = \sum_{i,j=1}^n a_i a_j C(\mathbf{s}_i, \mathbf{s}_j) \geq 0.$$

Es fundamental pues, desde el punto de vista metodológico, ofrecer garantías de que el modelo subyacente utilizado para explicar los datos es efectivamente un modelo válido. En este capítulo probaremos que existen funciones de correlación válidas en espacios no Euclídeos y repasaremos la literatura al respecto. De modo que pensamos que el problema de la validez de los modelos de segundo orden en problemas basados en el coste puede tener solución y vale la pena investigarlo.

Para eso será necesario analizar rigurosamente el problema teórico. Denotaremos por $D \subseteq \mathbb{R}^d$ al espacio de puntos donde el fenómeno en cuestión existe y tiene sentido. La estrategia será discutir las estructuras matemáticas con las que podemos dotar al conjunto D y, paralelamente, revisar la literatura existente sobre funciones definidas positivas en espacios no Euclídeos.

Ya hemos comprobado en la Propiedad 3.4 que el conjunto D con la distancia basada en el coste tiene estructura de espacio métrico. Más específicamente, en la Sección 5.1 veremos que D puede verse naturalmente como una Variedad Riemanniana, que tiene una estructura más abstracta y que induce naturalmente la estructura métrica discutida anteriormente. En el Apéndice E repasamos brevemente los conceptos básicos de la Geometría Riemanniana utilizados en el presente capítulo.

En la Figura 5.1 representamos a D en un esquema de relaciones entre estructuras matemáticas, que esperamos resulte de utilidad para seguir la discusión en este capítulo.

En dicho esquema, presentado como un grafo dirigido, las estructuras de mayor jerarquía son las más fundamentales y abstractas. De modo que cualquier nodo es un caso particular de sus *ancestros* y por tanto *hereda* toda su estructura y propiedades. Normalmente posee además alguna estructura o propiedad adicional. Así por ejemplo, todo espacio métrico es, en particular, un espacio premétrico en el que la función distancia verifica restricciones adicionales que transforman la *premétrica* en una *métrica*. O todo espacio de Banach es un espacio vectorial normado que además es completo.

Otra relación representada en el grafo con aristas destacadas en color es la de *inmersión* o *representación*. Como se muestra en el Apéndice E, toda

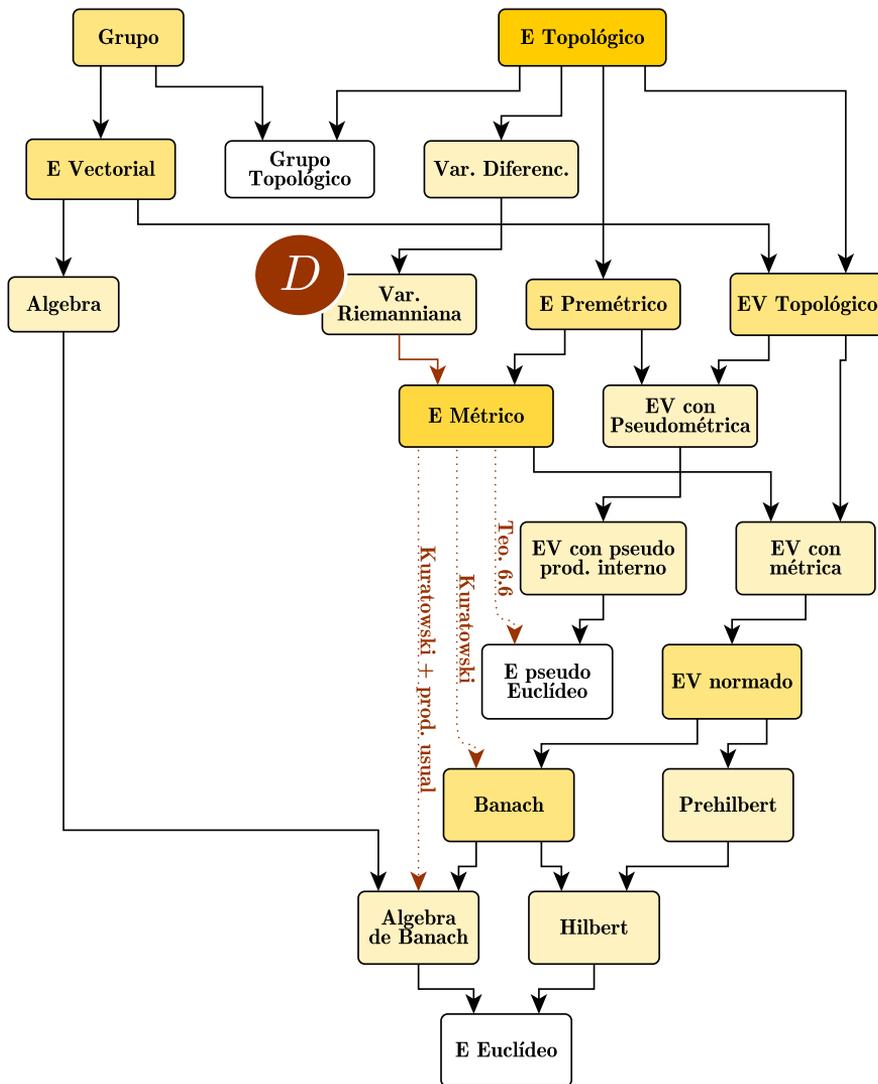


Figura 5.1. Ubicación del espacio D de puntos basado en el coste en una representación de la relación estructural entre diferentes tipos de espacios matemáticos.

Variedad Riemanniana induce un espacio métrico, si se define la función distancia adecuadamente. Así obtenemos un espacio más estructurado a partir de uno más abstracto y podemos ver a D bien como un espacio métrico, bien como una Variedad Riemanniana.

En el caso de las líneas discontinuas la inmersión es en un subconjunto de los espacios de llegada y la etiqueta de la arista indica el teorema que construye dicha inmersión. La inmersión de Kuratowski la describimos en este mismo capítulo, mientras que la inmersión en un espacio pseudo-Euclídeo la analizamos con detalle en el capítulo siguiente.

En la sección siguiente analizaremos las propiedades topológicas de D y desarrollaremos la estructura Riemanniana, comprobando que ésta induce la estructura métrica ya conocida.

A continuación veremos que D se puede representar en otros espacios con más estructura. Concretamente en un espacio de Banach, e incluso en un Álgebra de Banach, utilizando la inmersión de Kuratowski.

Finalmente, revisaremos la literatura existente y comentaremos las posibilidades de aplicar los respectivos resultados al caso que nos ocupa.

5.1 La Variedad Riemanniana basada en el coste

Típicamente, $D \subseteq \mathbb{R}^d$ es una región acotada de la superficie terrestre ($d = 2$, para regiones relativamente pequeñas, que se puedan aproximar por un plano), no necesariamente convexa, posiblemente con agujeros, aunque sí conexa. En algunas aplicaciones puede considerarse una región volumétrica ($d = 3$). Particularmente cuando se trabaja en la atmósfera o el mar.

Si fuera conveniente desde el punto de vista teórico, podríamos sin mayor dificultad considerar a D como una región no acotada, o sin agujeros. En este último caso bastaría *rellenar* los agujeros utilizando un coste muy alto en relación con el coste habitual en la región, tal como hicimos en la práctica para el cálculo de las distancias basadas en el coste.

Si contemplamos la posibilidad de existencia de agujeros, entonces D es un subconjunto del espacio vectorial \mathbb{R}^d , pero *no un subespacio vectorial* ya que, por ejemplo, no es cerrado para las operaciones. Si, en cambio, consideramos D como una región convexa, utilizando un coste infinito (o muy grande) para las barreras absolutas, es posible modelarlo como \mathbb{R}^d tal como se hace en la Geoestadística convencional. Sin embargo, esto resultaría inconveniente, ya que la estacionariedad e isotropía con respecto a la distancia en el coste no guarda ninguna relación con esta estructura vectorial. Podría decirse que la estructura vectorial del espacio Euclídeo es *incompatible* con la estacionariedad basada en el coste.

De la misma manera, tampoco es un *subespacio métrico* de \mathbb{R}^d , precisamente porque no nos interesa considerar la distancia Euclídea entre los puntos de D . En cambio, sí que es un *espacio métrico* con la distancia basada en el coste \mathfrak{d} inducida por una superficie de coste \mathfrak{f} , tal como se demuestra en la Propiedad 3.4.

La topología habitual T en \mathbb{R}^d induce naturalmente una estructura topológica restringida al subconjunto D

$$T_D = \{G \cap D : G \in T\}.$$

La estructura de espacio métrico con la distancia basada en el coste también induce una topología en D . Y además es la misma. Para verlo, basta razonar que todo abierto de T_D se puede incluir dentro de una bola abierta basada en el coste y viceversa. Por este motivo, la métrica basada en el coste y la inducida por la métrica Euclídea se dicen *equivalentes*.

Como consecuencia de ser espacio métrico, D cumple las propiedades topológicas de *1AN* (primer axioma de numerabilidad) y *Hausdorff*. Es decir, respectivamente, cada punto del espacio tiene una base de entornos numerable y puntos distintos tienen entornos disjuntos.

Otras propiedades topológicas que verifica el espacio D son:

- **Separabilidad** (i.e., tiene un subconjunto denso numerable). \mathbb{R}^d lo es, y todo subconjunto de un separable es separable. Por ejemplo, el conjunto

de vectores d -dimensionales con componentes racionales contenidos en D es un denso numerable de D .

- **2AN** (segundo axioma de numerabilidad, o *completamente separable*. Esto es, existe una base numerable de la topología). En un espacio métrico, la separabilidad equivale a 2AN.
- **Lindelöf**. Esto es una relajación de la compacidad. D no es compacto porque en principio no es cerrado. Pero todo cubrimiento abierto tiene un subcubrimiento numerable (en lugar de finito, como la compacidad). En un espacio métrico, separabilidad, 2AN y Lindelöf son equivalentes.

Nos interesa particularmente establecer si es posible que la función de distancia basada en el coste \mathfrak{d} resulte inducida por alguna estructura matemática más fundamental, como una norma o un producto interno. Esto dotaría al espacio de una estructura más rica y sofisticada, y nos proporcionaría más herramientas para el cálculo de las funciones definidas positivas.

La distancia \mathfrak{d} no puede provenir de una norma. Para empezar, una norma sólo se puede definir en un espacio vectorial, pero ya hemos visto que D no tiene esta estructura. Aunque consideráramos $D = \mathbb{R}^d$ con su estructura vectorial, esta función de distancia no podría provenir de una norma ya que, por ejemplo, no es invariante ante traslaciones. Esto es una manifestación de la incompatibilidad con la estructura vectorial Euclídea que mencionábamos al principio de esta sección.

En cambio, el espacio D puede considerarse *inmerso* en un espacio vectorial normado, cuya norma induzca esta distancia. La *inmersión de Kuratowski* (ver Sec. 5.2.1) permite identificar cualquier espacio métrico con un subconjunto del espacio de Banach $L^\infty(D)$. Esto es, el espacio vectorial normado completo de las funciones reales acotadas sobre D , con la norma supremo. Existen otras inmersiones posibles en espacios de Banach de dimensión infinita, como ℓ^∞ o $\mathcal{C}[0, 1]$ reseñadas en la Sección 5.2.

Incluso es posible dotar a este espacio con una estructura algebraica, utilizando el producto usual de funciones. El Álgebra de Banach resultante es conmutativa.

Más fundamentalmente, podemos ver a D como una *Variedad Riemanniana*, de modo tal que la métrica Riemanniana induzca la estructura de espacio métrico discutida en los párrafos precedentes.

En el caso de problemas con barreras absolutas (i.e., superficie de coste constante igual a 1, con *agujeros* de coste infinito) entonces se trata particularmente de una *subvariedad* de \mathbb{R}^2 de *codimensión* 0. La métrica Riemanniana Euclídea restringida a D induce el espacio métrico con la distancia basada en el coste.

En el caso general, $D \subseteq \mathbb{R}^2$ es una variedad con una métrica Riemanniana no Euclídea. Quizá el lector esté familiarizado con el Teorema de Nash ([Nash, 1956](#)), que garantiza que toda variedad Riemanniana se puede sumergir como subvariedad de un espacio Euclídeo de dimensión mayor. Pero en una subvariedad las geodésicas no tienen por qué ser líneas rectas, aunque la métrica Riemanniana sí que sea una métrica euclídea. Por tanto esta inmersión, aunque conserva la métrica Riemanniana, en general no conserva las distancias. De modo que no es una inmersión isométrica en tanto espacios métricos.

Una Métrica Riemanniana no debe confundirse con la métrica, o función de distancia, de un espacio métrico. En el Apéndice [E](#) definimos una métrica g en una variedad M como una familia de aplicaciones indexadas por M $\{g_p\}_{p \in M}$ tal que cada elemento g_p es un producto interno en el espacio tangente $T_p M$. Pero toda métrica Riemanniana induce una métrica τ_g en M . Ya que g determina la longitud de un vector tangente se puede definir la longitud de una curva $\alpha : [0, 1] \rightarrow M \in C^1$ como

$$L(\alpha) = \int_0^1 \sqrt{g_{\alpha(t)}(\alpha'(t), \alpha'(t))} dt$$

y definir la métrica τ_g como

$$\tau_g(s, t) = \inf_{\alpha \in D^1([0, 1]; M)_{(s, t)}} L(\alpha),$$

donde $D^1([0, 1]; M)_{(s, t)}$ es el conjunto de todas las aplicaciones C^1 a trozos $\alpha : [0, 1] \rightarrow M$ con $C(0) = s$, $C(1) = t$.

Notar que esta definición coincide con la distancia basada en el coste, siempre que $\sqrt{g_{\alpha(t)}(\alpha'(t), \alpha'(t))} dt$ represente el coste de desplazarse en dirección dt . Basta entonces con definir la métrica Riemanniana basada en el coste como

$$g_p(\mathbf{x}, \mathbf{y}) := \mathfrak{f}(p)^2 \langle \mathbf{x}, \mathbf{y} \rangle, \quad p \in D, \mathbf{x}, \mathbf{y} \in T_p D \quad (5.1)$$

donde \mathfrak{f} es la superficie de coste y $\langle \cdot, \cdot \rangle$ representa el producto interno euclídeo. De este modo, dada una curva α en D ,

$$L(\alpha) = \int_0^1 \sqrt{g_{\alpha(t)}(\alpha'(t), \alpha'(t))} dt = \int_0^1 \mathfrak{f}(\alpha(t)) \|\alpha'(t)\| dt.$$

Esto es, la longitud *Euclídea* de la curva ponderada en cada punto por el coste correspondiente. La métrica τ_g inducida por esta métrica Riemanniana es precisamente la distancia basada en el coste (Def. 3.3), donde el *coste* de una ruta α viene dado por la cantidad $L(\alpha)$.

5.2 Inmersiones en otros espacios

Definición 5.1 (Inmersión isométrica (isometric embedding) entre espacios métricos).

Dados dos espacios métricos X e Y , una inmersión isométrica es un homeomorfismo $f : X \rightarrow f(X) \in Y$ que preserva las distancias, i.e.,

$$d_X(a, b) = d_Y(f(a), f(b)), \quad \forall a, b \in X.$$

5.2.1 Inmersión de Kuratowski

El espacio métrico D se sumerge isométricamente en el espacio de Banach $L^\infty(D)$ formado por las funciones acotadas sobre D con la norma supremo

$$\|\phi\|_\infty := \sup_{x \in D} |\phi(x)|.$$

En efecto, fijado $x_0 \in D$, basta definir

$$\begin{aligned} D &\hookrightarrow L^\infty(D) \\ x &\mapsto \phi_x : D \rightarrow \mathbb{R} \\ & y \mapsto \mathfrak{d}(x, y) - \mathfrak{d}(y, x_0). \end{aligned}$$

Las funciones ϕ_x son efectivamente acotadas, ya que por la desigualdad triangular,

$$|\phi_x(y)| = |\mathfrak{d}(x, y) - \mathfrak{d}(y, x_0)| \leq \mathfrak{d}(x, x_0).$$

Y además, la norma induce una distancia en $L^\infty(D)$ que es compatible con la distancia basada en el coste definida en D .

$$\|\phi_{x_1} - \phi_{x_2}\|_\infty = \sup_{y \in D} |\phi_{x_1}(y) - \phi_{x_2}(y)| = \sup_{y \in D} |\mathfrak{d}(x_1, y) - \mathfrak{d}(x_2, y)| = \mathfrak{d}(x_1, x_2).$$

La última igualdad se consigue por la desigualdad triangular, que se hace igualdad en $y = x_1$ o $y = x_2$.

Esta norma no proviene de un producto interno. Si lo hiciera, habría de cumplir la geométrica *Regla del Paralelogramo* (ver Fig. 5.2): *En un paralelogramo, la suma de los cuadrados de sus lados es igual a la suma de los cuadrados de sus diagonales*. Y en este caso, si consideramos las funciones asociadas a los puntos x_1 y $x_2 \in D$ tenemos $\|\phi_{x_i}\|_\infty = \mathfrak{d}(x_i, x_0)$, $i = 1, 2$, por lo que

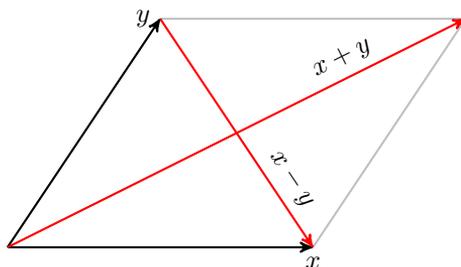
$$2(\|\phi_{x_1}\|_\infty^2 + \|\phi_{x_2}\|_\infty^2) = 2(\mathfrak{d}(x_1, x_0)^2 + \mathfrak{d}(x_2, x_0)^2).$$

Mientras que $\|\phi_{x_1} \pm \phi_{x_2}\|_\infty^2 = (\mathfrak{d}(x_1, x_0) \pm \mathfrak{d}(x_2, x_0))^2$, de modo que

$$\|\phi_{x_1} + \phi_{x_2}\|_\infty^2 + \|\phi_{x_1} - \phi_{x_2}\|_\infty^2 = 2(\mathfrak{d}(x_1, x_0) + \mathfrak{d}(x_2, x_0))^2,$$

por lo que la igualdad no se verifica.

Dejamos constancia de algunas ligeras variantes de la inmersión de Kuratowski, que en principio no afectan de ningún modo a nuestros propósitos.



$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$$

Figura 5.2. La regla del paralelogramo en un espacio pre-Hilbert.

Algunos autores consideran el conjunto más restringido de funciones reales *continuas* y acotadas, y otros el conjunto menos restringido de funciones esencialmente acotadas. En cualquier caso el espacio de llegada es un Banach.

Si el conjunto D es en sí mismo acotado (en las aplicaciones, habitualmente consideramos una región de trabajo acotada), la función asociada a un punto $x \in D$, ϕ_x , puede definirse más sencillamente como

$$\phi_x(y) = \mathfrak{d}(x, y).$$

En este caso, la función se corresponde con lo que llamábamos *mapa de distancias* (Def. 3.6) en el Capítulo 3.

5.2.2 Inmersión de Fréchet

Utilizando las propiedades de separabilidad del espacio métrico D , Heinson (2003) muestra que es posible sumergirlo isométricamente en el Banach ℓ^∞ de las sucesiones acotadas dotadas de la norma-supremo.

Basta para ello aplicar una inmersión de Kuratowski sobre un subconjunto numerable de D , teniendo en cuenta que $\ell^\infty = L^\infty(\mathbb{N})$.

Las ventajas teóricas son que el espacio de llegada es ahora independiente de D y además es de dimensión numerable. En principio, ninguna de estas ventajas supone diferencia alguna para nuestros propósitos.

5.2.3 Inmersión de Banach

También se puede mostrar (Heinonen, 2003, Teo. 3.6) que todo espacio métrico separable se sumerge isométricamente en el Banach $\mathcal{C}[0, 1]$ de las funciones reales continuas en el intervalo unidad cerrado, con la norma supremo.

5.3 Definición Positiva

La propiedad de definición positiva se enuncia de diferente manera según el espacio sobre el que se trabaje. Durante nuestra investigación, la naturaleza formal del espacio D está indeterminada. En la Sección 5.1 hemos visto que D se puede modelar como un espacio métrico, como una Variedad Riemanniana o incluso como un subconjunto de un Álgebra de Banach.

A continuación presentamos la definición positiva desde el punto de vista más abstracto posible y demostramos las propiedades fundamentales. Posteriormente repasamos la literatura sobre casos particulares donde el espacio base tiene estructuras más complejas, buscando puntos de intersección con el contexto de los problemas basados en el coste.

La siguiente definición se encuentra en Schoenberg (1938b), aunque nosotros la hemos modificado ligeramente quitando la simetría hermitica de la definición y deduciéndola como consecuencia.

Definición 5.2 (Funciones definidas positivas en espacios abstractos).¹

Sea \mathfrak{S} un conjunto cualquiera de elementos. Una función real o compleja F sobre $\mathfrak{S} \times \mathfrak{S}$ se dice *definida positiva*² si para cualesquiera $n \geq 2$ puntos

¹Moore (1935) incluye en la definición la simetría hermitica, que nosotros deducimos como propiedad.

²Algunos autores hablan de *semi-definida positiva* o *definida no-negativa* (e.g. Adler, 1981).

P_1, \dots, P_n en \mathfrak{S} y escalares $\alpha_1, \dots, \alpha_n$ reales o complejos,

$$\sum_{i,j=1}^n \alpha_i \bar{\alpha}_j F(P_i, P_j) \geq 0. \quad (5.2)$$

En particular, esta condición implica que el lado izquierdo de la desigualdad es un número real.

Si además \mathfrak{S} tiene estructura topológica, entonces también se requiere que la función $F(\cdot, \cdot)$ sea continua en ambas coordenadas. Denotamos a esta clase de funciones por $\mathcal{P}(\mathfrak{S})$.

Equivalentemente, F es definida positiva si para todo $n \geq 1$ y todo subconjunto $\{P_1, \dots, P_n\} \subseteq \mathfrak{S}$ las matrices $n \times n \left(F(P_i, P_j) \right)_{i,j=1}^n$ son semidefinidas positivas; y eventualmente, F es continua.

Propiedades 1. Si F es definida positiva en \mathfrak{S} , se demuestran las siguientes propiedades, que constituyen por tanto, condiciones *necesarias* de la definición positiva.

- $\forall P \in \mathfrak{S}$, $F(P, P)$ es un número real no negativo. En particular, la condición (5.2) se verifica desde $n = 1$.

Sean $n = 2$, $P_1 = P_2 = P$ y $\alpha_1 = \alpha_2 = \alpha$. Se tiene $\sum_{i,j=1}^2 \alpha_i \bar{\alpha}_j F(P_i, P_j) = 4F(P, P)|\alpha|^2 \geq 0$. Entonces,

$$0 \leq F(P, P) \in \mathbb{R}, \quad \forall P \in \mathfrak{S}. \quad (5.3)$$

En particular, si $n = 1$, $\sum_{i,j=1}^1 \alpha_i \bar{\alpha}_j F(P_i, P_j) = |\alpha|^2 F(P, P) \geq 0$.

- Simetría hermítica.

Sea $n = 2$. La condición de positividad (5.2) implica

$$\underbrace{|\alpha_1|^2 F(P, P) + |\alpha_2|^2 F(Q, Q) + \alpha_1 \bar{\alpha}_2 F(P, Q) + \alpha_2 \bar{\alpha}_1 F(Q, P)}_{\in \mathbb{R} \text{ (5.3)}} \in \mathbb{R}.$$

Por tanto,

$$\alpha_1 \bar{\alpha}_2 F(P, Q) + \overline{(\alpha_1 \bar{\alpha}_2 F(Q, P))} \in \mathbb{R}, \quad \forall \alpha_1, \alpha_2 \in \mathbb{C}.$$

Ahora, $\forall z_1, z_2, z \in \mathbb{C}$ $z_1 + \bar{z}_2 \in \mathbb{R} \Rightarrow \Im(z_1) = \Im(z_2)$, y $\Im(iz) = \Re(z)$. Por tanto,

$$\begin{aligned} \alpha_1 = \alpha_2 = 1 &\rightsquigarrow \Im(F(P, Q)) = \Im(\overline{F(Q, P)}) \\ \alpha_1 = i; \alpha_2 = 1 &\rightsquigarrow \Im(iF(P, Q)) = \Im(\overline{iF(Q, P)}) \\ &\Re(F(P, Q)) = \Re(\overline{F(Q, P)}), \end{aligned}$$

de donde,

$$F(P, Q) = \overline{F(Q, P)}, \quad \forall P, Q \in \mathfrak{S}. \quad (5.4)$$

- $\mathcal{P}(\mathfrak{S}) \neq \emptyset$. Por ejemplo, la función constante igual a 1 es definida positiva. $F(P, Q) \equiv 1$ es continua si el espacio es métrico y

$$\sum_{i,j=1}^n \alpha_i \bar{\alpha}_j F(P_i, P_j) = (\alpha_1 + \cdots + \alpha_n) \overline{(\alpha_1 + \cdots + \alpha_n)} = |(\alpha_1 + \cdots + \alpha_n)|^2 \geq 0.$$

Por tanto,

$$F(P, Q) \equiv 1 \in \mathcal{P}(\mathfrak{S}). \quad (5.5)$$

- $\Re(F(P, Q))$ es una cantidad acotada.

Sean $n = 2$, $P, Q \in \mathfrak{S}$. Por la condición de positividad y teniendo en cuenta la simetría hermítica se tiene, en particular,

$$\begin{aligned} \alpha_1 = \alpha_2 = 1 &\rightsquigarrow F(P, P) + F(P, Q) + F(Q, P) + F(Q, Q) = \\ &F(P, P) + F(Q, Q) + 2\Re(F(P, Q)) \geq 0 \\ \alpha_1 = -\alpha_2 = 1 &\rightsquigarrow F(P, P) - F(P, Q) - F(Q, P) + F(Q, Q) = \\ &F(P, P) + F(Q, Q) - 2\Re(F(P, Q)) \geq 0, \end{aligned}$$

de donde,

$$|\Re(F(P, Q))| \leq \frac{F(P, P) + F(Q, Q)}{2}. \quad (5.6)$$

- La clase $\mathcal{P}(\mathfrak{S})$ de funciones definidas positivas es un cono convexo.

Esto es, cualquier combinación lineal finita de funciones definidas positivas con coeficientes no negativos es definida positiva. Sean $\{F_i \in \mathcal{P}(\mathfrak{S})\}_{i=1}^m$, y $\{\lambda_i \geq 0\}_{i=1}^m$.

$$\sum_{i,j=1}^n \alpha_i \bar{\alpha}_j \left(\sum_{k=1}^m \lambda_k F(P_i, P_j) \right) = \sum_{k=1}^m \lambda_k \left(\sum_{i,j=1}^n \alpha_i \bar{\alpha}_j F(P_i, P_j) \right) \geq 0.$$

Así,

$$\forall \{F_i \in \mathcal{P}(\mathfrak{S})\}_{i=1}^m, \forall \{\lambda_i \geq 0\}_{i=1}^m, \sum_{k=1}^m \lambda_k F(P_i, P_j) \in \mathcal{P}(\mathfrak{S}).$$

- El producto de funciones definidas positivas es una función definida positiva.

Es una consecuencia de un lema de Schur, que establece que si dos formas cuadráticas $\sum_{i,k=1}^n a_{ik} \rho_i \rho_k$ y $\sum_{i,k=1}^n b_{ik} \rho_i \rho_k$ son positivas, entonces $\sum_{i,k=1}^n a_{ik} b_{ik} \rho_i \rho_k$ es también positiva (Schoenberg, 1938a). Por tanto,

$$\forall F_1, F_2 \in \mathcal{P}(\mathfrak{S}), \quad F_1 F_2(P_i, P_j) \in \mathcal{P}(\mathfrak{S}).$$

- Una función continua que es límite de una sucesión de funciones definidas positivas, es definida positiva.

Inmediato, por continuidad.

En esta Tesis hemos definido la definición positiva en el contexto más abstracto posible. Sin embargo, otros autores han utilizado definiciones alternativas utilizando características propias de los espacios concretos en los que han trabajado.

Un caso importante tiene lugar cuando \mathfrak{S} posee estructura de *Grupo*. La primera parte del libro de Sasvári (1994) está completamente dedicada a funciones definidas positivas sobre grupos y sus aplicaciones.³ Es particularmente destacable su versión del Teorema de Bochner (1.9.6) para grupos.

La diferencia cualitativa más importante que se obtiene al incorporar la estructura de grupo al conjunto $\mathfrak{S} = G$ es que entonces se puede hablar de funciones *estacionarias* con respecto a la operación del grupo (Adler y Taylor, 2007, Cap. 5). En efecto, asumiendo una notación aditiva, se dice que una función $F : G \times G \rightarrow \mathbb{R}$ es estacionaria si se puede escribir

$$F(p, q) = f(p - q), \quad \forall p, q \in G.$$

Así, por ejemplo, si f es tal que se puede escribir $f(p - q) = u(p)\overline{u(q)}$ entonces es definida positiva, ya que dados $p_1, \dots, p_n \in G$, la matriz cuyo elemento i, j viene dado por $f(p_i - p_j)$ se puede escribir

$$\begin{pmatrix} u(p_1) \\ \vdots \\ u(p_n) \end{pmatrix} \overline{\begin{pmatrix} u(p_1) & \dots & u(p_n) \end{pmatrix}} = \mathbf{u}\mathbf{u}^T,$$

que es claramente definida positiva. Bhatia (2007, p.144) utiliza este hecho para mostrar que la función $e^{i\langle \mathbf{a}, \mathbf{x} \rangle}$ es definida positiva en el espacio Euclídeo, $\forall \mathbf{a} \in \mathbb{R}^d$.

Si \mathfrak{S} tiene estructura de espacio vectorial, entonces la subclase de las funciones definidas positivas *estacionarias* está formada por aquellas que dependen únicamente del vector diferencia $F(\mathbf{x}_i, \mathbf{x}_j) = f(\mathbf{x}_j - \mathbf{x}_i)$. La función f queda entonces sujeta a las condiciones de continuidad y positividad

$$\sum_{i,j=1}^n \alpha_i \overline{\alpha_j} f(\mathbf{x}_i - \mathbf{x}_j) \geq 0.$$

³Muchos consideraríamos las *aplicaciones* de Sasvári como más bien teóricas, aunque de incuestionable utilidad para otras áreas de las matemáticas.

Para todo n , $\alpha_i \in \mathbb{C}$, y $\mathbf{x}_i \in \mathfrak{S}$, con $i = 1, \dots, n$. Esta es la definición más extendida en la literatura (véase, por ejemplo, [Schoenberg, 1938a](#)), particularmente con $\mathfrak{S} = E_m$, el espacio euclídeo de dimensión m , y está vinculada con las funciones de correlación de campos *estacionarios* en el espacio Euclídeo.

Si además \mathfrak{S} es un espacio vectorial *normado*, entonces una subclase de la anterior es aquella en la cual la dependencia es sólo sobre la longitud $\|\mathbf{x}_j - \mathbf{x}_i\|$ del vector:

$$F(\mathbf{x}_i, \mathbf{x}_j) = f(\mathbf{x}_j - \mathbf{x}_i) = g(\|\mathbf{x}_j - \mathbf{x}_i\|) = g(d(\mathbf{x}_i, \mathbf{x}_j)).$$

Aquí la función g es una función continua definida en $t \geq 0$ y $d(\cdot, \cdot)$ representa la *distancia* inducida por la norma.

Mientras que la función f es una función sobre un espacio vectorial, g está definida en $[0, \infty)$, dependiendo sólo de la estructura métrica del espacio. Así, se puede hacer una generalización, prescindiendo de la estructura vectorial de \mathfrak{S} y haciendo uso sólo de su estructura métrica (incluso semi-métrica; i.e., sin desigualdad triangular). Las funciones definidas positivas en un espacio semi-métrico, que sólo dependen de la distancia entre los elementos se pueden caracterizar como sigue. Esta caracterización se puede encontrar como definición en [Schoenberg \(1938b\)](#). Nosotros preferimos demostrar la equivalencia con la definición más abstracta de definición positiva.

Caracterización 5.3 (Funciones definidas positivas en espacios semi-métricos).

Sea \mathfrak{S} un espacio semi-métrico⁴ con función distancia $\mathfrak{d} : \mathfrak{S} \times \mathfrak{S} \rightarrow [0, +\infty)$. Las funciones definidas positivas que dependen únicamente de la distancia entre los elementos vienen dadas por las funciones reales y continuas g definidas sobre $\mathfrak{d}(\mathfrak{S} \times \mathfrak{S}) \subseteq [0, +\infty)$ tales que para cualesquiera $n \geq 2$ puntos P_1, \dots, P_n en \mathfrak{S}

⁴Espacio \mathfrak{S} provisto por una función distancia $\mathfrak{d} : \mathfrak{S} \times \mathfrak{S} \rightarrow [0, +\infty)$ simétrica que se anula sólo sobre la diagonal del producto cartesiano.

y escalares $\alpha_1, \dots, \alpha_n$ reales

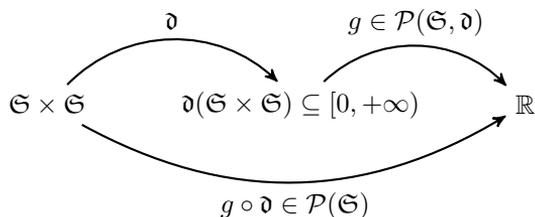
$$\sum_{i,j=1}^n \alpha_i \alpha_j g(\mathfrak{d}_{ij}) \geq 0, \tag{5.7}$$

donde $\mathfrak{d}_{ij} = \mathfrak{d}(P_i, P_j)$.

Si denotamos a esta clase de funciones por $\mathcal{P}(\mathfrak{S}, \mathfrak{d})$, podemos expresarlo sucintamente así: si \mathfrak{S} es un espacio semi-métrico,

$$g \in \mathcal{P}(\mathfrak{S}, \mathfrak{d}) \Leftrightarrow g \circ \mathfrak{d} \in \mathcal{P}(\mathfrak{S}). \tag{5.8}$$

Por extensión, a los elementos de $\mathcal{P}(\mathfrak{S}, \mathfrak{d})$ se les llama funciones definidas positivas en \mathfrak{S} .



Demostración.

(\Rightarrow) $F = g \circ \mathfrak{d}$ es composición de continuas y, por tanto, continua. Veamos que es definida positiva.

Sean $n \geq 2$, $P_1, \dots, P_n \in \mathfrak{S}$, $\alpha_1, \dots, \alpha_n \in \mathbb{C}$. Notar que la dificultad consiste en expresar la doble suma con coeficientes complejos en otra(s) con coeficientes reales, para lo cual bastan las identidades $z\bar{z} = |z|^2 = \Re(z)^2 + \Im(z)^2$; $z_i\bar{z}_j +$

$$\overline{(z_i \bar{z}_j)} = 2\Re(z_i \bar{z}_j) = 2\Re(z_i)\Re(z_j) + 2\Im(z_i)\Im(z_j).$$

$$\begin{aligned} \sum_{i,j=1}^n \alpha_i \bar{\alpha}_j F(P_i, P_j) &= \sum_{i=1}^n \alpha_i \bar{\alpha}_i g(\mathfrak{d}_{ii}) + \sum_{i,j=i+1}^n \left(\alpha_i \bar{\alpha}_j g(\mathfrak{d}_{ij}) + \alpha_j \bar{\alpha}_i g(\mathfrak{d}_{ji}) \right) \\ &= \sum_{i=1}^n |\alpha_i|^2 g(\mathfrak{d}_{ii}) + \sum_{i,j=i+1}^n (\alpha_i \bar{\alpha}_j + \overline{\alpha_i \bar{\alpha}_j}) g(\mathfrak{d}_{ij}) \\ &= \sum_{i,j=1}^n \Re(\alpha_i) \Re(\alpha_j) g(\mathfrak{d}_{ij}) + \sum_{i,j=1}^n \Im(\alpha_i) \Im(\alpha_j) g(\mathfrak{d}_{ij}) \geq 0. \end{aligned}$$

(\Leftarrow) Si F es definida positiva y depende sólo de la distancia entre los elementos, entonces se puede descomponer como $F = g \circ \mathfrak{d}$. Veamos que $g \in \mathcal{P}(\mathfrak{S}, \mathfrak{d})$. Dado $t \geq 0$, $\exists P, Q \in \mathfrak{S} / \mathfrak{d}(P, Q) = t$, puesto que g está definido sobre el recorrido de \mathfrak{d} .

1. g es real. Dado $t \in \mathfrak{d}(\mathfrak{S} \times \mathfrak{S})$ sean $P, Q \in \mathfrak{S} / \mathfrak{d}(P, Q) = t$,

$$g(t) = g(\mathfrak{d}(P, Q)) = g(\mathfrak{d}(Q, P)) \stackrel{[(5.4)]}{=} \overline{g(\mathfrak{d}(P, Q))} = \overline{g(t)}.$$

2. Positividad. Sea $n \geq 2$, $P_1, \dots, P_n \in \mathfrak{S}$, $\alpha_1, \dots, \alpha_n \in \mathbb{C}$,

$$\sum_{i,j=1}^n \alpha_i \bar{\alpha}_j g(\mathfrak{d}_{ij}) = \sum_{i,j=1}^n \alpha_i \bar{\alpha}_j F(P_i, P_j) \geq 0.$$

3. Continuidad. Si no lo fuera, sea t_0 un punto de discontinuidad y sean $P_0, Q_0 \in \mathfrak{S} / \mathfrak{d}(P_0, Q_0) = t_0$. Como \mathfrak{d} es continua, entonces $F = g \circ \mathfrak{d}$ sería discontinua en (P_0, Q_0) .

□

En el contexto de espacios semi-métricos, una función definida positiva g hereda algunas propiedades importantes, que constituyen condiciones *necesarias* de la definición positiva.

Propiedades 2.

- $\forall g \in \mathcal{P}(\mathfrak{S}, \mathfrak{d}), \quad g(0) \geq 0.$

Sea $P \in \mathfrak{S}$. $g(0) = g(\mathfrak{d}(P, P)) = F(P, P) \stackrel{(5.3)}{\geq} 0$. En particular, la Ecuación (5.7) se verifica desde $n = 1$.

- $B(\mathfrak{S}, \mathfrak{d}) \neq \emptyset.$

Sea $g(t) \equiv 1$. Pero $F = g \circ \mathfrak{d} \equiv 1 \stackrel{(5.5)}{\in} \mathcal{P}(\mathfrak{S})$. Por la Caracterización (5.8),

$$g(t) \equiv 1 \in \mathcal{P}(\mathfrak{S}, \mathfrak{d}). \quad (5.9)$$

- $g(t)$ es una cantidad acotada.

Sean $P, Q \in \mathfrak{S} / \mathfrak{d}(P, Q) = t$ y sea $F = g \circ \mathfrak{d} \stackrel{(5.6)}{\Rightarrow} |\Re(g(\mathfrak{d}(P, Q)))| \geq \frac{g(\mathfrak{d}(P, P)) + g(\mathfrak{d}(Q, Q))}{2}$. Pero g es real y $\mathfrak{d}(P, P) = \mathfrak{d}(Q, Q) = 0$ por propiedades de semi-métrica, entonces

$$|g(t)| \leq g(0).$$

5.4 Funciones definidas positivas en espacios no Euclídeos

El objetivo principal de la presente Tesis es aprender cuanto nos sea posible sobre la clase de funciones definidas positivas $\mathcal{P}(D)$.

En las secciones precedentes hemos visto que el espacio D es fundamentalmente una Variedad dotada con una métrica Riemanniana determinada por la superficie de coste \mathfrak{f} . Por otra parte, la estructura Riemanniana induce naturalmente una función de distancia: la distancia basada en el coste \mathfrak{d} que dota al espacio D de una estructura de espacio métrico.

Nos interesa particularmente pues, la subclase $\mathcal{P}(D, \mathfrak{d})$ de funciones definidas positivas *isotrópicas*, que dependen únicamente de la distancia basada en el coste.

En la presente sección haremos una revisión de la literatura existente acerca de las funciones definidas positivas en espacios no Euclídeos, examinando su vinculación con nuestro problema $\mathcal{P}(D, \mathfrak{d})$. También delinearemos posibles líneas de investigación en las que creemos que valdría la pena profundizar.

La primera pregunta que debemos responder es si existen funciones definidas positivas isotrópicas en espacios métricos no Euclídeos, más allá de la función trivial constante (5.9).

La respuesta es afirmativa, tal como lo demuestran resultados como los de Christakos y Papanicolaou (2000), o Curriero (2006). En ambos casos se considera el espacio \mathbb{R}^d dotado de las normas no Euclídeas $\|\cdot\|_p$ dadas por

$$\|\mathbf{x}\|_p = \left(|x_1|^p + \cdots + |x_d|^p \right)^{1/p}.$$

En este contexto prueban que las funciones de correlación isotrópicas de la forma

$$f(\mathbf{h}) = e^{-\|\mathbf{h}\|_p^\beta},$$

para ciertos valores de los parámetros p y β son definidas positivas. En particular, esto prueba que la función de correlación exponencial ($p = 1$) es válida con la distancia Manhattan.

Por lo tanto, existen funciones de covarianza no triviales que son válidas para algunas métricas no Euclídeas.

Por otra parte, Christakos et al. (2000) comprueban la definición positiva de diversas funciones calculando numéricamente las densidades espectrales (i.e., resolviendo las integrales utilizando métodos de cuadratura).

El interesantísimo trabajo de Christakos, muy bien detallado y ampliado en su libro (Christakos, 2000), nos enseña, entre otras cosas, que existen familias de funciones definidas positivas isotrópicas con respecto a distancias no Euclídeas.

Aunque los ejemplos concretos que utilizan son limitados y necesitan de un análisis de validez *ad-hoc*.

Sin embargo, estos resultados no son directamente aplicables a problemas basados en el coste. Fundamentalmente porque la distancia basada en el coste no puede provenir de una norma, como hemos establecido en la Sección 5.1. Además, el método numérico de cálculo de la densidad espectral sólo es aplicable en espacios de dimensión finita y provistos de una estructura prehilbertiana.

Podrían aplicarse si fuéramos capaces de sumergir isométricamente D en \mathbb{R}^d con una norma de esta familia. Pero todas las inmersiones que hemos analizado en la Sección 5.2 son en espacios de dimensión infinita con la norma supremo. Concretamente en espacios (o álgebras) de Banach.

En el contexto de álgebras de Banach, una de las generalizaciones del Teorema de Bochner más abstractas que hemos encontrado es la de Rudin (1991, Teo. 11.32) (junto con la de Sasvári (1994, Teo. 1.9.6) para grupos). Esta es una línea de trabajo futuro que consideramos muy interesante y prometedora, que incluso podría realizarse en colaboración con un especialista en la Teoría Espectral de álgebras de Banach.

Schoenberg (1938b), además del Teorema 2.5 de caracterización de funciones definidas positivas isotrópicas del espacio Euclídeo E_d , también hizo un aporte significativo en la abstracción del resultado. En la Sección 2 habla sobre funciones definidas positivas en un espacio de Hilbert. Allí, el Teorema 2 se puede replantear en los siguientes términos:

Si H es un espacio de Hilbert de dimensión infinita y $\Phi : H \rightarrow \mathbb{R}$ es una función definida positiva continua en el grupo H que depende sólo de la norma, i.e., $\Phi(x) = \varphi(\|x\|)$ para cierta $\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}$ continua, entonces $\varphi(\sqrt{t})$ es la transformada de Laplace de alguna medida finita μ sobre \mathbb{R}_+ .

Aunque no tenemos estructura Hilbertiana en el espacio D , ni tenemos inmersión en un Hilbert, Berg et al. (1984, Teo. 5§2.1) llevan este resultado al

contexto más abstracto de los espacios de Banach y muestra que si $\varphi(\|x\|)$ es definida positiva en un espacio de Banach, entonces $\varphi(\sqrt{t})$ es la transformada de Laplace de alguna medida finita sobre \mathbb{R}_+ .

Hemos visto en la Sección 5.2 que es posible sumergir D en un subconjunto de un Banach, por lo que este resultado podría ser de interés. Sin embargo, Berg et al. (1984, Teo. 5§2.3) establecen que todas las funciones definidas positivas del Banach $\mathcal{C}[0, 1]$ son las constantes no negativas. Es decir, las funciones trivialmente definidas positivas que ya conocemos.

Naturalmente, puesto que D se representa en un *subconjunto* del Banach $\mathcal{C}[0, 1]$, sólo podemos concluir que

$$\mathcal{P}(\mathcal{C}[0, 1]) \subseteq \mathcal{P}(D, \mathfrak{d}).$$

Por tanto, que $\mathcal{P}(\mathcal{C}[0, 1])$ sea trivial, no significa que $\mathcal{P}(D, \mathfrak{d})$ también lo sea.

Finalmente, hemos visto que la definición positiva se puede estudiar en el contexto de un Grupo, donde el concepto de estacionariedad adquiere su carácter más abstracto.

La representación espectral de un campo aleatorio sobre un grupo está íntimamente relacionado con la Teoría de Representación del grupo. Es difícil proporcionar una teoría general, ya que la representación se hace en términos del grupo de *caracteres* de G , que obviamente depende de G . En efecto, en el grupo implícito en el espacio Euclídeo \mathbb{R}^d la familia de caracteres son precisamente las funciones exponenciales complejas que aparecen en la transformada de Fourier del Teorema de Bochner.

Desde el punto de vista teórico, otras referencias fundamentales son Rudin (1962) acerca del Análisis de Fourier sobre grupos, o más recientemente Berg et al. (1984) sobre Análisis Armónico sobre semi-grupos.

En cualquier caso el espacio D no tiene, en principio, estructura de grupo. Sin embargo, existe una generalización también mencionada en Adler y Taylor (2007, §5.3) y más detallada en Bochner (1941) en la cual consideramos los elementos g de un grupo G *actuando* sobre los elementos P del espacio subyacente \mathfrak{S} .

La *acción* de un elemento del grupo sobre un punto es una abstracción de la aplicación *traslación* $\mathbf{x} \mapsto \mathbf{x} + \mathbf{h}$ en un espacio Euclídeo. Así, para todo $g \in G$, el elemento $gP \in \mathfrak{S}$ es el elemento *transformado* de P por la acción de g .

De este modo, se pueden considerar funciones *invariantes con respecto al grupo*,

$$F(gP, gQ) = F(P, Q), \quad \forall g \in G, \forall P, Q \in \mathfrak{S},$$

lo que es una abstracción de las funciones de *estacionarias* del espacio Euclídeo. Bochner (1941, §III) obtiene una expansión ortogonal de las funciones sobre $\mathfrak{S} \times \mathfrak{S}$, invariantes respecto a un grupo que actúa sobre \mathfrak{S} , en términos de una base de funciones continuas, tal y como en una descomposición espectral. Y consigue además una caracterización de los coeficientes de la expansión, de modo tal que la función sea definida positiva. La base de funciones depende del grupo y del espacio concretos.

Estos resultados nos parecen de particular relevancia, porque muestran que es posible obtener una caracterización de las funciones definidas positivas estacionarias en contextos mucho más abstractos que el espacio Euclídeo. Una línea abierta de investigación es la búsqueda de un grupo de *movimientos rígidos* que actúe sobre el espacio D , que nos permita aprovechar estos resultados. El grupo natural para este caso es el implícito en el espacio vectorial \mathbb{R}^d , actuando de modo que si $\mathbf{x} \in \mathbb{R}^d$ y $P \in D$, el punto $\mathbf{x}P \in D$ sea el que resulte de recorrer la geodésica que pasa por P en dirección \mathbf{x} a lo largo de una distancia (basada en el coste) $\|\mathbf{x}\|$. Seguir esta línea implica un trabajo de verificación de ciertas hipótesis asumidas por Bochner (1941), identificar la base de funciones ortogonales para producir finalmente funciones definidas positivas estacionarias utilizando la caracterización de Bochner. Aún restaría utilizar la métrica en D para encontrar funciones isotrópicas, que es en última instancia lo que nos interesa fundamentalmente.

5.5 Conclusiones

En este capítulo analizamos rigurosamente el problema matemático de la búsqueda de funciones definidas positivas isotrópicas en problemas basados en el coste. Una familia de funciones que denotamos por $\mathcal{P}(D, \mathfrak{d})$.

En primer lugar estudiamos las propiedades topológicas del problema y acabamos por ver que su naturaleza fundamental se puede describir como una variedad Riemanniana. La métrica Riemanniana (Ec. 5.1) se define naturalmente a partir de la superficie de coste asociada al problema, e induce precisamente la distancia basada en el coste, con la estructura métrica que conlleva.

La literatura sobre definición positiva en espacios no Euclídeos, sin embargo, siempre hace uso de estructuras matemáticas adicionales, como estructuras de Grupo o vectorial. Por tanto, una estrategia ha sido analizar la posibilidad de sumergir el espacio D en algún otro espacio más estructurado.

En la Proposición 6.4 veremos que no es posible, en general, sumergir el espacio D en un espacio Euclídeo, ni en un Hilbert de dimensión infinita. Sin embargo, en la Sección 5.2 hemos visto cómo se puede sumergir D en un Álgebra de Banach utilizando la inmersión de Kuratowski. Por otra parte, en el Capítulo 6 analizaremos en detalle la inmersión en un espacio pseudo-Euclídeo.

Siendo tan variado el espectro de espacios donde nos podemos plantear el problema, fue necesario hacer una revisión de la definición positiva en diferentes contextos. Planteamos la definición en un conjunto abstracto, discutimos sus propiedades y, posteriormente, revisamos las definiciones más concretas en espacios más estructurados.

Por último, repasamos los resultados conocidos sobre definición positiva en espacios no Euclídeos y comentamos su vinculación con el problema $\mathcal{P}(D, \mathfrak{d})$.

Empezamos por asegurar que existen soluciones no triviales de problemas no Euclídeos, aunque no sean de aplicación directa a nuestro problema.

La línea de trabajo que pasa por la inmersión de D en un espacio de Banach cuenta con la interesante generalización del Teorema de Bochner sobre álgebras

de Banach (Rudin, 1991, Teo. 11.32). Tenemos además la caracterización de Berg et al. (1984, Teo. 5§2.1) sobre funciones definidas positivas isotrópicas en espacios de Banach. Aunque también sabemos (Berg et al., 1984, Teo. 5§2.3) que éstas son triviales en el caso del Banach $\mathcal{C}[0, 1]$, precisamente uno de los espacios donde sabemos representar D .

Éste es un camino aún por recorrer, del cual los teoremas y resultados comentados son hitos que es imprescindible tener en cuenta y aprovechar. En su caso, es necesario un estudio teórico más riguroso de la Teoría Espectral en álgebras de Banach que el que hemos esbozado aquí.

Finalmente, y evitando las inmersiones, podemos considerar el grupo de movimientos rígidos del espacio vectorial \mathbb{R}^d actuando sobre la variedad D . Una *traslación* en este contexto se efectuaría recorriendo la geodésica que pasa por el punto de origen, a lo largo de una distancia determinada por la longitud del vector traslación. Podemos utilizar resultados en Bochner (1941) para encontrar una representación espectral de funciones definidas positivas estacionarias. Por otra parte, la representación espectral en una variedad implica una definición de la integración con respecto a una medida adecuada. Una posibilidad natural podría ser la medida de Hausdorff (Federer, 1978), que generaliza la de Lebesgue en \mathbb{R}^d , y permite así considerar una transformación de Fourier generalizada (Lindgren et al., 2011). La caracterización de la isotropía, sin embargo, queda por explorar. En este particular, es interesante el trabajo de Strichartz (1974) sobre las transformadas de Fourier de funciones invariantes ante la acción de un grupo (de rotaciones).

Aún así, consideramos que ésta es la línea de trabajo futuro más *pura* y promisoria.

Capítulo 6

Representación pseudo-Euclídea

En el capítulo anterior se ha visto que una estrategia interesante para tratar la definición positiva en espacios no Euclídeos consiste en la inmersión de la variedad Riemanniana (o el espacio métrico que ésta induce) D en otro espacio con mayor estructura. En este capítulo veremos, sin embargo, que no es posible sumergirlo en un espacio Euclídeo (Prop. 6.4). Si así fuera, podríamos utilizar inmediatamente las herramientas geoestadísticas habituales en dicha inmersión. Sin embargo es posible hacerlo de forma aproximada utilizando la técnica conocida como Escalado Multidimensional (Multidimensional Scaling; MDS), introducida en la Sección 6.1. Aunque, como discutiremos en dicha sección, es una solución que no resulta satisfactoria.

En cambio, sí es posible sumergir D de forma exacta en un espacio pseudo-Euclídeo, siguiendo las ideas del MDS y relajando las restricciones adecuadas. En este capítulo introducimos las ideas necesarias del MDS que utilizaremos para la construcción de nuestra representación pseudo-Euclídea. Definiremos rigurosamente el espacio y sus propiedades y veremos que cualquier problema

basado en el coste se puede representar en él. Además, la caracterización de las funciones definidas positivas dada por el Teorema de Bochner sigue siendo válida en el espacio pseudo-Euclídeo.

6.1 Multidimensional Scaling

Bajo el nombre genérico de MDS se agrupan una colección de técnicas para el análisis de datos de similitud (o disimilitud) en un conjunto de objetos. Estos datos pueden ser las correlaciones entre las respuestas de una encuesta, índices de similitud entre candidatos políticos, etc. Aunque fue originalmente desarrollado por psicometricistas, ha devenido en una técnica estándar en el análisis de datos multivariantes y es ampliamente utilizado en diversas disciplinas.

Nos referiremos aquí al llamado *Classical Scaling* (Torgerson, 1952, 1958; Gower, 1966; Cuadras et al., 1998). Una visión en profundidad de todas las variantes y aplicaciones del MDS se puede encontrar en Borg y Groenen (2005).

El MDS pretende construir configuraciones de puntos de un espacio geométrico (euclidiano) a partir de la información sobre la disimilitud entre los elementos de un conjunto, de modo que las distancias entre los puntos representen las correspondientes disimilitudes. La motivación habitual de esta técnica es obtener una representación gráfica de la estructura de los datos, que es mucho más fácil de entender que una tabla de números y además sintetiza la información esencial contenida en los datos, suavizando el ruido.

En nuestro caso, sin embargo, la motivación es otra. No buscamos una representación gráfica de nuestros datos, ya que éstos son esencialmente geográficos. Lo que nos interesa es la capacidad de la técnica MDS de generar una aproximación euclidiana de nuestros datos: las distancias basadas en el coste entre localizaciones. Porque de este modo, en un espacio Euclídeo, se pueden utilizar los resultados ya conocidos de la Teoría Geoestadística.

Pero como se dijo, esta aproximación se hace a costa de *suavizar el ruido*. Lo que en determinadas aplicaciones puede considerarse *ruido*, en nuestro contexto

constituye una componente importante de lo que pretendemos explicar. Cuando utilizamos distancias no Euclídeas como la basada en el coste, es porque queremos capturar un fenómeno que no se puede explicar con distancias Euclídeas y lo último que nos interesa es *filtrarlo*.

Antes de proceder a discutir los detalles de implementación, podemos visualizar estos conceptos en acción si aplicamos el MDS al conjunto de distancias basadas en el coste del ejemplo piloto.

Hemos encontrado que con un espacio Euclídeo de 2 dimensiones conseguimos explicar aproximadamente el 80 % de la distribución espacial. Es ilustrativo comparar el comportamiento de los tres tipos de distancias. En la Figura 4.6 se puede ver para cada pareja de puntos en este juego de observaciones (en orden creciente de distancia basada en el coste) la distancia euclídea y la aproximación MDS con respecto a la distancia basada en el coste.

Se pueden observar, desde otro punto de vista, los mismos efectos que en la Figura 4.5. Por ejemplo, el incremento notable de la variabilidad a partir de los 95 m, que también se observa en la aproximación MDS. O también la divergencia progresiva entre las distancias Euclídea y basada en el coste.

Se ve claramente que el efecto principal de la aproximación MDS es el de corregir esta divergencia, aproximando las correspondientes distancias basadas en el coste de una manera que podríamos llamar *insesgada*. Por otra parte, la variabilidad en el error de aproximación se mantiene prácticamente igual que la de la distancia Euclídea.

En este sentido, podemos concluir que la aproximación MDS determina una configuración euclidiana que corrige la desviación media, a largo plazo, pero no la variabilidad a pequeña escala que es la que produce las incongruencias que han justificado el desarrollo de la distancia basada en el coste.

De todas maneras haremos a continuación una introducción técnica al MDS, porque nos servirá de base para el desarrollo de la Representación pseudo-Euclídea, que es una generalización.

6.1.1 Disimilitudes y distancias

Definición 6.1. Disimilitud y matriz de disimilitudes

Consideramos las *disimilitudes* entre n objetos como un conjunto

$$\{\delta_{ij} \in \mathbb{R}; i, j = 1, \dots, n\}$$

de números reales que verifica los axiomas básicos de *reflexividad* ($\delta_{ii} = 0$), *simetría* ($\delta_{ij} = \delta_{ji}$) y *no-negatividad* ($\delta_{ij} \geq 0$) $\forall i, j = 1, \dots, n$.

Resulta conveniente recoger los opuestos de los semi-cuadrados de estas cantidades en una matriz

$$\mathbf{\Delta} = \left(-\frac{1}{2}\delta_{ij}^2\right)_{n \times n}$$

a la que nos referiremos como la *matriz de disimilitudes*.

Notar que la matriz es simétrica, con componentes menores o iguales que cero y tiene diagonal nula, como consecuencia de los axiomas de las disimilitudes.

Las *distancias* entre los puntos de un espacio métrico verifican esta definición. Cabe recordar que habíamos establecido en el Capítulo 5 que un conjunto de localizaciones junto con la distancia basada en el coste conformaban un espacio métrico.

Aunque las distancias verifican además la condición más restrictiva de la *desigualdad triangular*, son aún suficientemente flexibles como para modelar un amplio abanico de situaciones. Por ejemplo, en Teoría de Grafos, los *pesos* de las aristas son usualmente no-negativos y métricos (i.e., satisfacen la desigualdad triangular), pero no necesariamente euclidianos.

Gower y Legendre (1986) proporcionan un ejemplo sencillo de un grafo métrico no-euclidiano que reproducimos en la Figura 6.1. Los puntos P_1, P_2, P_3 forman un triángulo equilátero de lado 2, y el punto P_4 es equidistante (1,1 unidades a P_1, P_2 y P_3). En un espacio Euclídeo, la distancia mínima a la que P_4 puede estar de los otros vértices es $2/\sqrt{3} \approx 1,15$. Por tanto este conjunto de distancias no se puede representar en un espacio Euclídeo de ninguna dimensión.

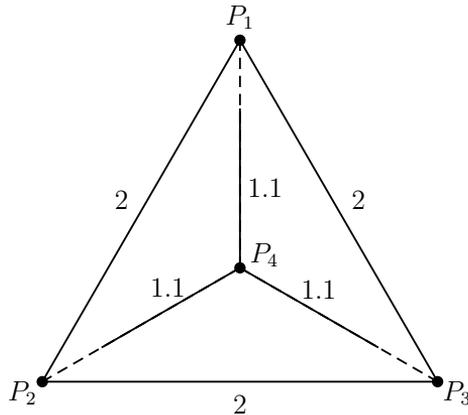


Figura 6.1. Un conjunto de puntos en un espacio métrico que no tienen una representación Euclídea.

Definición 6.2. Matrices Euclídeas

Cuando se puede encontrar en algún espacio Euclídeo un conjunto de n puntos tal que las $n(n-1)/2$ distancias dos a dos generan los valores δ_{ij} , se dice que la matriz $\mathbf{\Delta}$ es Euclídea.

[Gower \(1985\)](#) desarrolla las propiedades de las matrices Euclídeas, en particular, la condición para su caracterización.

6.1.2 Classical Scaling

Asumamos que existe un conjunto de puntos $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ en un espacio Euclídeo E_d de dimensión desconocida d , cuyas distancias dos a dos se corresponden exactamente con las disimilitudes de partida, es decir

$$\begin{aligned} \delta_{ij}^2 &= (\mathbf{x}_j - \mathbf{x}_i)'(\mathbf{x}_j - \mathbf{x}_i) \\ -\frac{1}{2}\delta_{ij}^2 &= \mathbf{x}_i'\mathbf{x}_j - \frac{1}{2}\mathbf{x}_i'\mathbf{x}_i - \frac{1}{2}\mathbf{x}_j'\mathbf{x}_j \quad i, j = 1, \dots, n. \end{aligned} \quad (6.1)$$

Obtener una representación euclidiana exacta de las distancias codificadas en la matriz $\mathbf{\Delta}$ equivale a resolver el sistema de ecuaciones anterior.

Notar que estas ecuaciones implican necesariamente que $\delta_{ii} = 0$, $i = 1, \dots, n$ y que $\delta_{ij} = \delta_{ji}$, $i, j = 1, \dots, n$. Por tanto, puesto que $\mathbf{\Delta}$ es una matriz simétrica con diagonal nula, (6.1) es un sistema de $n(n-1)/2$ ecuaciones con nd incógnitas.

Por otra parte, d debe ser menor o igual que $n-1$, ya que no tiene sentido considerar un conjunto de n puntos en un espacio de dimensión mayor: siempre se puede encontrar una rotación tal que las últimas coordenadas de todos los puntos sean nulas. Esto es la generalización del hecho conocido de que en un espacio Euclídeo, dos puntos ($n=2$) determinan una recta ($d=1$), tres puntos ($n=3$) determinan un plano ($d=2$), etc. Siempre puede ocurrir que los tres puntos estén alineados y se puedan incluir en una recta. O, en general, es posible que un conjunto de n puntos pueda ser descrito en un espacio de dimensión menor que $n-1$. En este caso se habla de un conjunto *degenerado*.

Excepto en casos muy degenerados ($d = (n-1)/2$), el sistema es indeterminado. Gower (1982) establece la condición necesaria y suficiente para que el sistema sea compatible, o dicho de otra manera, para que la matriz $\mathbf{\Delta}$ sea Euclídea. Procederemos constructivamente, asumiendo que la solución existe y derivando el resultado. En el proceso obtendremos las condiciones necesarias y finalmente veremos que son suficientes.

Sea $\mathbf{X}_{n \times d}$ la matriz de coordenadas de los puntos, de modo que la fila i de \mathbf{X} viene dada por el vector \mathbf{x}_i . El Sistema (6.1) puede reformularse en términos de la siguiente ecuación matricial

$$\mathbf{\Delta} = \mathbf{X}\mathbf{X}' - \frac{1}{2}\text{diag}(\mathbf{X}\mathbf{X}')\mathbf{1}' - \frac{1}{2}\mathbf{1}\text{diag}(\mathbf{X}\mathbf{X}')', \quad (6.2)$$

donde $\text{diag}(\mathbf{A})$ representa la diagonal de la matriz \mathbf{A} como un vector columna, y $\mathbf{1} = (1, \overset{(n)}{\cdot}, 1)'$.

Notar que si \mathbf{X} es una solución válida, entonces cualquier *movimiento* geométrico producirá la misma estructura de distancias y por tanto, otra solución válida. Así que necesitamos imponer restricciones adicionales a \mathbf{X} para encontrar una fórmula para una solución única.

Una restricción respecto de la traslación puede ser localizar alguno de los puntos \mathbf{x}_i en el origen de coordenadas. La alternativa estándar en MDS consiste en situar en el origen al baricentro o centroide del conjunto de puntos. Es inmediato *centrar* una matriz de coordenadas \mathbf{X} cualquiera: basta multiplicar por la izquierda por la *matriz centradora* $\mathbf{H} = \mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}'$. Notar que $\mathbf{H}\mathbf{X} = \mathbf{X} - \mathbf{1}(\frac{1}{n}\mathbf{1}'\mathbf{X})$, por tanto este producto consiste en restar a cada elemento la media de la columna correspondiente. Una matriz de coordenadas \mathbf{X} está centrada si $\mathbf{H}\mathbf{X} = \mathbf{X}$ y entonces verifica que $\mathbf{1}'\mathbf{X} = \mathbf{1}'\mathbf{H}\mathbf{X} = \mathbf{0}$, ya que la matriz \mathbf{H} tiene la propiedad $\mathbf{1}'\mathbf{H} = \mathbf{H}\mathbf{1} = \mathbf{0}$. Esta propiedad nos permite simplificar la Ecuación (6.2). Multiplicando por izquierda y derecha por \mathbf{H} , y asumiendo que la matriz \mathbf{X} está centrada:

$$\mathbf{B} := \mathbf{H}\Delta\mathbf{H} = \mathbf{X}\mathbf{X}' \quad (6.3)$$

La matriz $\mathbf{X}\mathbf{X}'$ es la llamada *matriz de Gram* de los vectores $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ y es simétrica y semidefinida positiva.

Esta ecuación implica que la matriz \mathbf{B} de disimilitudes *doblemente centrada* debe ser *necesariamente* semidefinida positiva para que exista una solución. Veremos además que esta condición es *suficiente*, ya que, asegurada esta condición, seremos capaces de calcular una solución \mathbf{X} de la Ecuación (6.3).

El sistema sigue estando indeterminado respecto de la rotación del sistema de coordenadas, por lo que si pretendemos calcular una solución unívoca del sistema, debemos fijar la orientación. Esta indeterminación se resuelve de forma estándar utilizando un sistema de coordenadas principales. \mathbf{B} es una matriz real y simétrica y por tanto admite una descomposición en valores y vectores propios, con valores propios *reales* y no-negativos. Así, se puede descomponer de forma unívoca como

$$\mathbf{B} = \mathbf{\Gamma}\mathbf{\Lambda}\mathbf{\Gamma}' = \mathbf{X}\mathbf{X}'$$

donde $\mathbf{\Gamma}$ es la matriz de vectores propios unitarios y $\mathbf{\Lambda}$ es la matriz diagonal de valores propios reales no-negativos, ordenados en forma decreciente.

Notar que en esta descomposición de \mathbf{B} se puede prescindir de los vectores propios asociados al valor propio 0 y de los valores propios correspondientes.

Es seguro que \mathbf{B} tiene valor propio 0 pues como mínimo el vector $\mathbf{1}$ está en su núcleo, por propiedades de la matriz \mathbf{H} : $\mathbf{B}\mathbf{1} = \mathbf{H}\mathbf{\Delta}(\mathbf{H}\mathbf{1}) = 0$. Por tanto $\mathbf{\Gamma}$ es una matriz $n \times d$, con $d \leq n - 1$.

Una solución del sistema es, por tanto,

$$\mathbf{X} = \mathbf{\Gamma}\mathbf{\Lambda}^{1/2},$$

que existe siempre que \mathbf{B} sea real y semidefinida positiva.

La matriz \mathbf{X} es entonces una matriz $n \times d$, con $d \leq n - 1$, cuyas columnas son los vectores propios de la matriz $\mathbf{B} = \mathbf{H}\mathbf{\Delta}\mathbf{H}$ normalizados según la magnitud del valor propio correspondiente y cuyas filas proporcionan las coordenadas de los puntos $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$.

Queda demostrada así la caracterización de las matrices Euclidianas, que se puede resumir en el siguiente teorema.

Teorema 6.3. *Caracterización de las matrices de disimilitudes Euclidianas*

Una matriz de disimilitudes $\mathbf{\Delta}$ es Euclidiana si y sólo si

$$\mathbf{B} := \mathbf{H}\mathbf{\Delta}\mathbf{H} \geq 0.$$

Y una representación Euclidiana viene dada por la fórmula

$$\mathbf{X} = \mathbf{\Gamma}\mathbf{\Lambda}^{1/2},$$

donde $\mathbf{\Gamma}$ es la matriz $n \times d$ de vectores propios unitarios con valores propios no nulos y la matriz $\mathbf{\Lambda}$ es la matriz $d \times d$ diagonal de los correspondientes valores propios.

Como corolario, notar que el número de valores propios nulos ($\dim \ker \mathbf{B} \geq 1$) proporciona una medida de la *degeneración* del conjunto de puntos.

En la práctica, el *Classical Scaling* utiliza este teorema para obtener representaciones Euclidianas *aproximadas* de estructuras de distancias que no son

necesariamente Euclidianas, o como estrategia para reducir la dimensionalidad de un problema. Para ello se toman los primeros (típicamente 2) valores y vectores propios (positivos) de la descomposición espectral de \mathbf{B} para la construcción de una solución aproximada \mathbf{X} . Naturalmente, los puntos obtenidos en la matriz \mathbf{X} no tienen por qué subtender exactamente las distancias codificadas en Δ , pero constituyen la mejor aproximación posible en la medida que se consideran las direcciones propias de mayor magnitud.

El cociente entre la suma de estos valores propios y la suma del total de valores propios proporciona una medida de la variabilidad espacial explicada por la aproximación considerada.

6.2 Representación pseudo-Euclídea

Ya hemos justificado en la Sección 6.1 que no nos interesa una representación Euclídea aproximada, ya que esto *filtraría* las particularidades espaciales de pequeña escala que son precisamente las que queremos capturar utilizando la distancia basada en el coste.

Nos preguntamos entonces, si es posible obtener una representación Euclídea *exacta* de una estructura basada en el coste, aunque sea en un espacio de mayor dimensión, incluso infinita (i.e., en un espacio de Hilbert). Para responder a esta pregunta, necesitamos plantear el problema en términos de una inmersión entre espacios matemáticos (Def. 5.1, p.102).

Hemos establecido en el Capítulo 5 que un conjunto de localizaciones espaciales junto con la distancia basada en el coste conformaban un espacio métrico que denotamos por D . Por tanto, desde el punto de vista matemático, la representación Euclídea exacta puede verse como un problema de inmersión isométrica entre el espacio métrico D y el espacio Euclídeo E_n , que es otro espacio métrico.

El siguiente resultado muestra que no es posible, en términos generales, obtener una representación Euclídea exacta de un problema de distancias basadas en el coste.

Proposición 6.4 (Inmersibilidad isométrica en el espacio Euclídeo E_n o en el Hilbert \mathcal{H}).

*El espacio métrico D con la distancia basada en el coste **no es**, en general, sumergible isométricamente en un espacio Euclídeo de ninguna dimensión, ni tampoco en un espacio de Hilbert \mathcal{H} (de dimensión infinita).*

Demostración.

Basta proporcionar un caso para el cual no se pueda encontrar una representación Euclídea exacta en ninguna dimensión finita o infinita.

Utilizaremos el ejemplo de [Curriero \(2006\)](#) que introdujimos en la Sección 3.2.3. Aquí tenemos cuatro puntos en el plano, con matriz de distancias

$$\mathbf{D} = \begin{pmatrix} 0 & 1 & 1 & 2 \\ 1 & 0 & 2 & 1 \\ 1 & 2 & 0 & 1 \\ 2 & 1 & 1 & 0 \end{pmatrix}.$$

Evidentemente no hay ninguna configuración de cuatro puntos en el plano euclídeo con esa matriz de distancias. Pero es que tampoco existe tal configuración en E_3 ni en ningún espacio euclídeo E_n , $n \geq 2$. Esto se comprueba fácilmente con la caracterización establecida en el Teorema 6.3. La matriz $\mathbf{B} := \mathbf{H}\mathbf{D}\mathbf{H}$, con $\mathbf{\Delta} = (-\frac{1}{2}d_{ij}^2)$ tiene valores propios $\{2, 2, 0, -1\}$, por lo que no es semidefinida positiva.

Así, la matriz del ejemplo de Curriero *no es Euclidiana* y por tanto no existe ninguna representación exacta de estas distancias en ningún espacio euclídeo, por grande que sea su dimensión.

Teniendo en cuenta que el espacio métrico D es separable (ver Cap. 5), para el caso de dimensión infinita podemos aplicar el siguiente Teorema de Menger, referenciado en [Schoenberg \(1937, p.788\)](#):

Un espacio métrico D puede ser inmerso isométricamente en el Hilbert \mathfrak{H} si y sólo si D es separable y todo conjunto de $n + 1$ ($n \geq 2$) puntos distintos de D pueden ser inmersos en E_n .

Así pues, nuestro espacio D *tampoco es sumergible en un Hilbert*, ya que los cuatro puntos del ejemplo de Curriero no pueden sumergirse en E_3 . \square

Puesto que el espacio Euclídeo queda descartado, podríamos preguntarnos si existe algún otro producto interno (i.e., una forma bilineal simétrica y definida positiva), distinto del Euclídeo, que permita representar nuestro espacio D en un \mathbb{R}^n dotado con ése producto (i.e., en un espacio pre-Hilbert).

La respuesta es que no. Si lo hubiera, podríamos representar el producto con la matriz \mathbf{S} tal que $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}'\mathbf{S}\mathbf{y}$ y, siguiendo el razonamiento de la Sección 6.1.2, la Ecuación (6.3) se transformaría en

$$\mathbf{H}\mathbf{\Delta}\mathbf{H} = \mathbf{X}\mathbf{S}\mathbf{X}',$$

para cierta matriz \mathbf{X} , cuyas filas son los vectores que representan los puntos que subtienden las disimilitudes en $\mathbf{\Delta}$.

Como un producto interno es definido positivo, entonces \mathbf{S} lo es y también $\mathbf{X}\mathbf{S}\mathbf{X}'$. Y ya hemos comprobado, con el ejemplo de Curriero, que la matriz $\mathbf{B} = \mathbf{H}\mathbf{\Delta}\mathbf{H}$ tiene valores propios negativos, por lo que la ecuación no tiene solución, para ningún producto interno \mathbf{S} .

De modo que necesitamos flexibilizar aún más la estructura del espacio de representación. Relajando la condición de definición positiva en el producto interno, obtenemos un espacio donde sí es posible representar el espacio métrico D . Este es el espacio pseudo-Euclídeo.

6.2.1 Inmersibilidad en un espacio pseudo-Euclídeo

Hemos establecido que el espacio métrico D no es sumergible isométricamente en ningún espacio Euclídeo, ni en un espacio de Hilbert y tampoco en un pre-Hilbert, que son sucesivas generalizaciones del primero.

Demostraremos a continuación que sí es sumergible en un espacio pseudo-Euclídeo, que es una generalización aún mayor.

Definición 6.5 (Espacio pseudo-Euclídeo).

Es un espacio vectorial real de dimensión finita d , junto con un producto interno pseudo-Euclídeo de la forma $(\mathbf{x}, \mathbf{y}) = (x_1y_1 + \cdots + x_ky_k) - (x_{k+1}y_{k+1} + \cdots + x_dy_d)$. Esto es, una forma bilineal no-degenerada y simétrica $(,) : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$, que verifica

$$\text{Simetría : } (\mathbf{x}, \mathbf{y}) = (\mathbf{y}, \mathbf{x}), \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^d$$

$$\text{Bilinealidad : } (\lambda \mathbf{x} + \mathbf{y}, \mathbf{z}) = \lambda(\mathbf{x}, \mathbf{z}) + (\mathbf{y}, \mathbf{z}), \quad \forall \lambda \in \mathbb{R}, \forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^d$$

$$\text{No-degeneración : } (\mathbf{x}, \mathbf{y}) = 0 \quad \forall \mathbf{y} \in \mathbb{R}^d \Rightarrow \mathbf{x} = \mathbf{0}.$$

Notar que es como un producto interno, excepto porque la propiedad de definición positiva ha sido reemplazada por la condición más débil de no-degeneración.

El número d se llama la *dimensión* del espacio, mientras que k se llama el *índice*. El par $(k, d - k)$ se llama la *signatura*. El espacio pseudo-Euclídeo de dimensión d e índice k se denota por $E_{(k, d-k)}$. El producto de dos vectores \mathbf{x} , \mathbf{y} del espacio $E_{(k, d-k)}$ lo expresaremos en términos de la *matriz de signatura* $\mathbf{S}_k = \text{diag}(1, \dots, 1, -1, \dots, -1)$ y el producto usual de matrices:

$$(\mathbf{x}, \mathbf{y}) = \mathbf{x}' \mathbf{S}_k \mathbf{y}.$$

Este producto pseudo-Euclídeo induce una *premétrica simétrica* en el espacio $E_{(k, d-k)}$ dada por

$$d(\mathbf{x}, \mathbf{y}) := \sqrt{|(\mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y})|}. \quad (6.4)$$

Esto es una función real en $\mathbb{R}^d \times \mathbb{R}^d$, con las propiedades de *no-negatividad* ($d(\mathbf{x}, \mathbf{y}) \geq 0$) y *simetría* ($d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$); sin embargo no hay *identidad de indiscernibles* ($d(\mathbf{x}, \mathbf{y}) = 0 \iff \mathbf{x} = \mathbf{y}$) ni tampoco *desigualdad triangular* ($d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$).¹ En cambio, sí se cumple que $d(\mathbf{x}, \mathbf{x}) = 0 \quad \forall \mathbf{x}$.

¹Esencialmente porque no se cumple la Desigualdad de Cauchy-Schwartz ($|(\mathbf{x}, \mathbf{y})|^2 \leq (\mathbf{x}, \mathbf{x})(\mathbf{y}, \mathbf{y})$) por la existencia de vectores $\mathbf{y} \neq \mathbf{0}$ que sin embargo $(\mathbf{y}, \mathbf{y}) = 0$.

Podemos hacernos una imagen mental intuitiva de cómo funciona este peculiar espacio considerando el espacio pseudo-Euclídeo de dos dimensiones con índice 1, $E_{(1,1)}$ y observando el ejemplo de la Figura 6.2.

Los vectores afines al eje horizontal (que varían a lo largo de la primera dimensión) tienen cuadrado positivo ($\overline{OA}^2 = (\overline{OA}, \overline{OA}) = \overline{OC}^2 = 1$) como en el caso Euclídeo, mientras que los que varían a lo largo de la segunda dimensión tienen cuadrado negativo ($\overline{OB}^2 = -1$). Concretamente, el opuesto del cuadrado de la distancia Euclídea correspondiente. En general, el cuadrado de un vector será una combinación de las componentes de cada dimensión. Específicamente, los puntos situados a lo largo de las diagonales distan cero entre sí. En particular, vemos cómo se viola la desigualdad triangular puesto que $2 = d(A, C) \not\leq d(A, B) + d(B, C) = 0$.

Notar que, como la forma bilineal no es definida positiva, es necesario el valor absoluto en la definición de la premétrica. Como hemos comprobado, existen puntos en $E_{(k,d-k)}$ de *cuadrado negativo* $(\mathbf{x}, \mathbf{x}) < 0$. En el espacio pseudo-Euclídeo existe un *cono* formado por los puntos de *cuadrado nulo* y que divide el espacio en una región de puntos de cuadrado negativo y otra de puntos de cuadrado positivo. Al cuadrado de la diferencia entre dos puntos $(\mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y})$ le llamaremos *distancia cuadrática con signo*, teniendo en cuenta que puede ser una cantidad positiva o negativa y que, por tanto, puede diferir en signo con el cuadrado de la distancia $d(x, y)^2$.

Siguiendo los pasos del *Classical Scaling*, consideremos otra vez la matriz de disimilitudes $\mathbf{\Delta} = \left(-\frac{1}{2}\delta_{ij}^2\right)$, que recoge las distancias $\delta_{ij} \geq 0$ dos a dos en un conjunto de n puntos $\{P_1, \dots, P_n\} \subseteq D$. Supongamos que existe un conjunto de puntos $\{\mathbf{x}_1, \dots, \mathbf{x}_n\} \subseteq E_{(k,d-k)}$ que serán las representaciones respectivas de

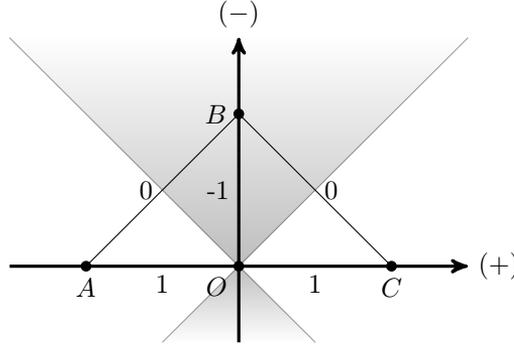


Figura 6.2. Algunos puntos y sus distancias relativas en el espacio pseudo-Euclídeo $E_{(1,1)}$

P_1, \dots, P_n en el espacio pseudo-Euclídeo y tales que *conserven las distancias*:²

$$\begin{aligned}\delta_{ij}^2 &= (\mathbf{x}_j - \mathbf{x}_i)' \mathbf{S}_k (\mathbf{x}_j - \mathbf{x}_i) \\ &= \mathbf{x}_i' \mathbf{S}_k \mathbf{x}_i + \mathbf{x}_j' \mathbf{S}_k \mathbf{x}_j - 2\mathbf{x}_i' \mathbf{S}_k \mathbf{x}_j \\ -\frac{1}{2}\delta_{ij}^2 &= \mathbf{x}_i' \mathbf{S}_k \mathbf{x}_j - \frac{1}{2}\mathbf{x}_i' \mathbf{S}_k \mathbf{x}_i - \frac{1}{2}\mathbf{x}_j' \mathbf{S}_k \mathbf{x}_j.\end{aligned}$$

Sea \mathbf{X} la matriz $n \times d$ cuyas filas vienen dadas por las coordenadas de los puntos $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$. Entonces el elemento (i, j) de la matriz $n \times n$ $\mathbf{X}\mathbf{S}_k\mathbf{X}'$ es precisamente $\mathbf{x}_i' \mathbf{S}_k \mathbf{x}_j$. Así, la matriz \mathbf{X} debe verificar

$$\mathbf{\Delta} = \mathbf{X}\mathbf{S}_k\mathbf{X}' - \frac{1}{2}\text{diag}(\mathbf{X}\mathbf{S}_k\mathbf{X}')\mathbf{1}' - \frac{1}{2}\mathbf{1}\text{diag}(\mathbf{X}\mathbf{S}_k\mathbf{X}'),$$

donde $\text{diag}(\mathbf{A})$ denota la diagonal de la matriz \mathbf{A} dispuesto en forma de columna y $\mathbf{1} = (1, \dots, 1)'$.

Notar la analogía con la Ecuación (6.2). Por los mismos argumentos mencionados entonces, estamos frente a un sistema indeterminado ante traslaciones y rotaciones. Para encontrar una solución única procederemos de la misma manera,

²En el sentido de que la distancia cuadrática con signo coincida con los cuadrados de las distancias preestablecidas.

asumiendo que la matriz \mathbf{X} está centrada y multiplicando la ecuación por derecha e izquierda por la matriz centradora \mathbf{H}

$$\mathbf{B} := \mathbf{H}\mathbf{\Delta}\mathbf{H} = \mathbf{X}\mathbf{S}_k\mathbf{X}'.$$

Finalmente, para eliminar la indeterminación ante rotaciones, utilizamos la descomposición espectral de la matriz \mathbf{B} :

$$\mathbf{B} = \mathbf{\Gamma}\mathbf{\Lambda}\mathbf{\Gamma}' = \mathbf{X}\mathbf{S}_k\mathbf{X}',$$

donde $\mathbf{\Gamma}$ es la matriz de vectores propios unitarios y $\mathbf{\Lambda}$ es la matriz diagonal de valores propios reales, ordenados en forma decreciente.

A diferencia del *Classical Scaling* esta ecuación tiene solución *siempre*, no exigiendo que \mathbf{B} sea definida positiva gracias a la presencia de la matriz de signatura.

Igual que en el caso clásico, en esta descomposición de \mathbf{B} se puede prescindir de los vectores propios asociados al valor propio 0, entre ellos, como mínimo el $\mathbf{1}$.

El sistema se puede resolver entonces reescribiendo $\mathbf{\Lambda} = (\mathbf{\Lambda}\mathbf{S}_k)^{1/2}\mathbf{S}_k(\mathbf{S}_k\mathbf{\Lambda}')^{1/2}$ —lo cual determina el índice k del espacio pseudo-Euclídeo como el número de valores propios no-negativos (positivos) de la matriz \mathbf{B} —, pues entonces $\mathbf{B} = \left[\mathbf{\Gamma}(\mathbf{\Lambda}\mathbf{S}_k)^{1/2}\right]\mathbf{S}_k\left[\mathbf{\Gamma}(\mathbf{\Lambda}\mathbf{S}_k)^{1/2}\right]' = \mathbf{X}\mathbf{S}_k\mathbf{X}'$ y la matriz \mathbf{X} se obtiene por identificación

$$\mathbf{X} = \mathbf{\Gamma}(\mathbf{\Lambda}\mathbf{S}_k)^{1/2}.$$

La matriz \mathbf{X} es entonces una matriz $n \times d$, con $d \leq n - 1$, cuyas columnas son los vectores propios (menos los del núcleo) de la matriz $\mathbf{B} = \mathbf{H}\mathbf{\Delta}\mathbf{H}$ normalizados según la magnitud del valor propio correspondiente y cuyas filas proporcionan las coordenadas de los puntos $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$.

Notar que la matriz $\mathbf{\Lambda}\mathbf{S}_k$ es sencillamente la matriz diagonal de los módulos de los valores propios, ya que \mathbf{S}_k es precisamente la matriz diagonal de signos de los valores propios.

Hemos sentado así las bases para el siguiente teorema.

Teorema 6.6 (Representación de un conjunto de puntos en un espacio pseudo-Euclídeo).

Dado el conjunto de distancias dos a dos $\{\delta_{ij}\}_{i,j=1}^n$, con $\delta_{ji} = \delta_{ij} \geq 0$ y $\delta_{ii} = 0$, no necesariamente Euclídeas entre n elementos $\{P_1, \dots, P_n\}$, siempre es posible encontrar un conjunto de puntos $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ en el espacio pseudo-Euclídeo $E_{(k,d-k)}$ que conserven las distancias.

Concretamente, las coordenadas de una solución (no única) vienen dadas por las filas de la matriz

$$\mathbf{X} = \mathbf{\Gamma}(\mathbf{\Lambda}\mathbf{S}_k)^{1/2}, \quad (6.5)$$

obtenida a partir de la descomposición espectral de la matriz $\mathbf{H}\mathbf{\Delta}\mathbf{H} = \mathbf{\Gamma}\mathbf{\Lambda}\mathbf{\Gamma}'$ con $\mathbf{\Delta} = (-\frac{1}{2}\delta_{ij}^2)$ y la matriz de signatura \mathbf{S}_k determinada por los signos de los valores propios.

Demostración.

Como $\mathbf{\Delta}$ y \mathbf{H} son simétricas, la matriz $\mathbf{B} = \mathbf{H}\mathbf{\Delta}\mathbf{H}$ es simétrica y, por el Teorema Espectral, tiene una descomposición $\mathbf{\Gamma}\mathbf{\Lambda}\mathbf{\Gamma}'$ con valores propios reales y n vectores propios ortogonales reales. Puesto que 0 es valor propio con vector propio $\mathbf{1}$, podemos excluirlos de la descomposición, junto con otros vectores propios que pudiera haber asociados al valor propio 0. Así, $\mathbf{\Lambda}$ es una matriz diagonal $d \times d$, con $d \leq n - 1$, con k valores propios positivos y $d - k$ valores propios negativos, ordenados en forma descendente. Además asumiremos que los vectores propios en $\mathbf{\Gamma}$ son unitarios.

Construimos la *matriz de signatura* \mathbf{S}_k diagonal con los k primeros elementos iguales a 1 y los $d - k$ restantes iguales a -1 , de modo que $(\mathbf{\Lambda}\mathbf{S}_k)^{1/2}$ es simplemente la matriz diagonal de las raíces cuadradas de los valores absolutos de los valores propios.

Por este procedimiento siempre se puede construir la matriz \mathbf{X} de tamaño $n \times d$, con la Fórmula (6.5).

Ahora veamos que las filas de esta matriz subtienden las distancias pseudo-Euclídeas correspondientes. La disimilitud δ_{ij}^2 entre los puntos \mathbf{x}_i y \mathbf{x}_j es el elemento (i, j) de la matriz

$$\text{diag}(\mathbf{X}\mathbf{S}_k\mathbf{X}')\mathbf{1}' + \mathbf{1}\text{diag}(\mathbf{X}\mathbf{S}_k\mathbf{X}')' - 2\mathbf{X}\mathbf{S}_k\mathbf{X}'.$$

Pero por construcción, $\mathbf{X}\mathbf{S}_k\mathbf{X}' = \mathbf{\Gamma}\mathbf{\Lambda}\mathbf{\Gamma}' = \mathbf{H}\mathbf{\Delta}\mathbf{H}$, y además,

$$\begin{aligned}\mathbf{H}\mathbf{\Delta}\mathbf{H} &= \left(\mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}'\right)\mathbf{\Delta}\left(\mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}'\right) \\ &= \mathbf{\Delta} - \frac{1}{n}\mathbf{1}\mathbf{1}'\mathbf{\Delta} - \frac{1}{n}\mathbf{\Delta}\mathbf{1}\mathbf{1}' + \frac{1}{n^2}\mathbf{1}\mathbf{1}'\mathbf{\Delta}\mathbf{1}\mathbf{1}'.\end{aligned}$$

Puesto que $\frac{1}{n}\mathbf{1}'\mathbf{\Delta}$, $\frac{1}{n}\mathbf{\Delta}\mathbf{1}$ y $\frac{1}{n^2}\mathbf{1}'\mathbf{\Delta}\mathbf{1}$ son las medias por columnas $-\frac{1}{2}\delta_{\cdot j}^2$, filas $-\frac{1}{2}\delta_{i\cdot}^2$ y global $-\frac{1}{2}\delta_{\cdot\cdot}^2$ de la matriz $\mathbf{\Delta}$ respectivamente, el elemento (i, j) de la matriz $\mathbf{H}\mathbf{\Delta}\mathbf{H}$ es $-\frac{1}{2}(\delta_{ij}^2 - \delta_{\cdot j}^2 - \delta_{i\cdot}^2 + \delta_{\cdot\cdot}^2)$. Teniendo en cuenta que la diagonal de la matriz $\mathbf{\Delta}$ es nula y que por su simetría las medias por filas y por columnas son iguales, se tiene

$$\begin{aligned}(\mathbf{x}_i - \mathbf{x}_j, \mathbf{x}_i - \mathbf{x}_j) &= -\frac{1}{2}(\delta_{ii}^2 - \delta_{i\cdot}^2 - \delta_{\cdot i}^2 + \delta_{\cdot\cdot}^2) - \\ &\quad \frac{1}{2}(\delta_{jj}^2 - \delta_{\cdot j}^2 - \delta_{j\cdot}^2 + \delta_{\cdot\cdot}^2) + (\delta_{ij}^2 - \delta_{\cdot j}^2 - \delta_{i\cdot}^2 + \delta_{\cdot\cdot}^2) = \delta_{ij}^2.\end{aligned}$$

En particular, el signo de la distancia cuadrática entre \mathbf{x}_i y \mathbf{x}_j es positivo y se conservan las distancias ($d(\mathbf{x}_i, \mathbf{x}_j) = \delta_{ij}$). \square

Sólo tiempo después de desarrollar este teorema encontramos que ya había sido utilizado en el campo informático de reconocimiento de patrones ([Peřkalska y Duin, 2005](#); [Goldfarb, 1985](#)).

6.2.2 Fórmulas adicionales

La Fórmula (6.5) permite obtener inmediatamente la representación pseudo-Euclídea de cualquier conjunto de disimilitudes. En la práctica sin embargo, es

frecuente la necesidad de añadir puntos adicionales luego de un cierto procesamiento (ver, por ejemplo, Gower, 1968 en el caso Euclídeo).

Por ejemplo, se usa un conjunto de puntos (observaciones) para estimar los parámetros de un modelo espacial y luego se usan estas estimaciones para predecir en otro conjunto de puntos distinto. Pero no se puede volver a utilizar la Fórmula (6.5) para obtener las coordenadas pseudo-Euclídeas de este nuevo conjunto de puntos. Para empezar, es necesario que estén en el mismo sistema de referencia que los primeros, por lo que habría que transformar los puntos de observación y predicción simultáneamente, previo al análisis. Si bien esto es posible, es conveniente tener la posibilidad de añadir puntos de predicción *a posteriori*.

Por tanto, es útil proporcionar fórmulas que permitan añadir puntos a una representación ya calculada, manteniendo las coordenadas de los puntos originales y respetando las disimilitudes entre los puntos nuevos y los originales, y también entre sí mismos.

Partimos entonces de una matriz \mathbf{X} $n \times d$ cuyas filas son las coordenadas de un conjunto de puntos $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^d$ que subtienden dos a dos las disimilitudes simétricas y no-negativas δ_{ij} recogidas en la matriz $\mathbf{\Delta} = (-\frac{1}{2}\delta_{ij}^2)$, con diagonal nula.

Ahora se tiene un nuevo punto P del cual se conocen las disimilitudes con cada uno de los n puntos anteriores, recogidas en el vector $\boldsymbol{\delta} = (\delta_i^2) = (d(P, P_i)^2)$. Se pretende encontrar las coordenadas del punto P en el mismo sistema de coordenadas que el resto de los puntos.

En general, el conjunto de $n + 1$ puntos tendrá *dimensionalidad* n . Si \mathbf{X} es no degenerada, entonces $d = n - 1$ y será necesario considerar los n vectores originales $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^d \subset \mathbb{R}^{d+1} = \mathbb{R}^n$ añadiendo una componente adicional nula. Así, el punto P tendrá coordenadas $\mathbf{x}_P \equiv (x_{P1}, \dots, x_{Pd}, x_{P(d+1)})'$. En cambio, si \mathbf{X} es degenerada, entonces puede ser necesario añadir más de una dimensión nueva, con coordenadas no nulas para los puntos originales.

Ocasionalmente el punto P se encontrará en el mismo hiperplano de \mathbb{R}^{d+1} que los vectores originales y no será necesario añadir la nueva dimensión. Como

se verá más adelante, este caso quedará naturalmente incluido en las fórmulas que se obtienen para la situación más general.

Así, en el caso en que \mathbf{X} es no degenerada, el problema se reduce al sistema de n ecuaciones con $d + 1 \leq n$ incógnitas determinado por las distancias entre \mathbf{x}_P y cada uno de los n vectores originales. Descomponiendo convenientemente $\mathbf{x}_P = (\mathbf{x}', x_{P(d+1)})'$,

$$\begin{aligned} \delta_i^2 &= (\mathbf{x}_i - \mathbf{x})' \mathbf{S}_k (\mathbf{x}_i - \mathbf{x}) \pm x_{P(d+1)}^2, \\ &= \mathbf{x}'_i \mathbf{S}_k \mathbf{x}_i + \mathbf{x}' \mathbf{S}_k \mathbf{x} - 2\mathbf{x}'_i \mathbf{S}_k \mathbf{x} \pm x_{P(d+1)}^2, \quad i = 1, \dots, n, \end{aligned} \quad (6.6)$$

donde el signo del último término se elige para que el sistema siempre tenga solución y determina el índice del espacio pseudo-Euclídeo de dimensión $d + 1$ en el que se representan los $n + 1$ puntos. Notar que este índice es independiente del sistema de coordenadas elegido, ya que depende del número de valores propios positivos de la matriz \mathbf{B} , como ha quedado establecido en la sección anterior.

Conviene representar el Sistema de ecuaciones (6.6) de forma matricial. Recordar de la sección anterior que la matriz \mathbf{X} verifica la relación

$$\mathbf{X} \mathbf{S}_k \mathbf{X}' = \mathbf{H} \mathbf{\Delta} \mathbf{H} \equiv \mathbf{B};$$

por tanto, $\mathbf{x}'_i \mathbf{S}_k \mathbf{x}_i$ se puede expresar en términos de los elementos de la diagonal de la matriz \mathbf{B} , que son independientes del sistema de coordenadas elegido. Así, el Sistema de ecuaciones (6.6) es equivalente a la ecuación matricial

$$\boldsymbol{\delta} = \text{diag}(\mathbf{B}) + (\mathbf{x}' \mathbf{S}_k \mathbf{x} \pm x_{P(d+1)}^2) \mathbf{1} - 2\mathbf{X} \mathbf{S}_k \mathbf{x}. \quad (6.7)$$

Notar que el número real $\mathbf{x}'\mathbf{S}_k\mathbf{x} \pm x_{P(d+1)}^2$ puede verse como la *norma cuadrática*³ del vector incógnita \mathbf{x}_P , en virtud de su descomposición:

$$\|\mathbf{x}_P\|^2 = \mathbf{x}'_P \begin{pmatrix} \mathbf{S}_k & 0 \\ 0 & \pm 1 \end{pmatrix} \mathbf{x}_P = \mathbf{x}'\mathbf{S}_k\mathbf{x} \pm x_{P(d+1)}^2,$$

donde $\begin{pmatrix} \mathbf{S}_k & 0 \\ 0 & \pm 1 \end{pmatrix}$ constituye la nueva matriz de signatura (aún indeterminada) en el espacio pseudo-Euclídeo de dimensión $d + 1$. Este número puede calcularse en términos de las distancias cuadráticas $\boldsymbol{\delta}$ y de la diagonal de \mathbf{B} multiplicando la Ecuación (6.7) por el vector $\mathbf{1}'$, es decir, sumando todas las ecuaciones del Sistema (6.6):

$$\begin{aligned} \mathbf{1}'\boldsymbol{\delta} &= \mathbf{1}'\text{diag}(\mathbf{B}) + \|\mathbf{x}_P\|^2\mathbf{1}'\mathbf{1} - 2\underbrace{\mathbf{1}'\mathbf{X}\mathbf{S}_k\mathbf{x}}_{=0} \\ &= \mathbf{1}'\text{diag}(\mathbf{B}) + n\|\mathbf{x}_P\|^2, \end{aligned}$$

así,

$$\begin{aligned} \|\mathbf{x}_P\|^2 &\equiv \mathbf{x}'\mathbf{S}_k\mathbf{x} \pm x_{P(d+1)}^2 = \frac{1}{n}\mathbf{1}'(\boldsymbol{\delta} - \text{diag}(\mathbf{B})) \\ &= -\frac{1}{n}\mathbf{1}'\mathbf{d}, \end{aligned} \tag{6.8}$$

con $\mathbf{d} = \text{diag}(\mathbf{B}) - \boldsymbol{\delta}$. Ahora, sustituyendo en (6.7) y reordenando,

$$\begin{aligned} 2\mathbf{X}\mathbf{S}_k\mathbf{x} &= (\text{diag}(\mathbf{B}) - \boldsymbol{\delta}) - \frac{1}{n}\mathbf{1}\mathbf{1}'(\text{diag}(\mathbf{B}) - \boldsymbol{\delta}) \\ &= \mathbf{H}\mathbf{d}, \end{aligned}$$

Premultiplicando por \mathbf{X}' y teniendo en cuenta que como \mathbf{X} es centrada $\mathbf{X}'\mathbf{H} = \mathbf{X}'$, se tiene,

$$\mathbf{x} = \frac{1}{2}(\mathbf{X}'\mathbf{X}\mathbf{S}_k)^{-1}\mathbf{X}'\mathbf{d}. \tag{6.9}$$

³Aunque no se ha definido la norma en este espacio, considérese como una notación cómoda para el cuadrado $(\mathbf{x}_P, \mathbf{x}_P)$ del vector.

Notar que la matriz $\mathbf{X}'\mathbf{X}\mathbf{S}_k$ es efectivamente invertible pues $\mathbf{X}'\mathbf{X}$ lo es por ser definida positiva y \mathbf{S}_k también lo es ya que $\mathbf{S}_k\mathbf{S}_k = \mathbf{I}$. Por tanto, ella misma es su inversa y $(\mathbf{X}'\mathbf{X}\mathbf{S}_k)^{-1} = \mathbf{S}_k(\mathbf{X}'\mathbf{X})^{-1}$.

En el caso particular en el cual los vectores $\mathbf{x}_1, \dots, \mathbf{x}_n$ están en coordenadas principales —i.e., la matriz \mathbf{X} fue calculada como en la sección anterior, por la Ecuación (6.5)—, entonces

$$\begin{aligned}\mathbf{X}'\mathbf{X}\mathbf{S}_k &= \left[\boldsymbol{\Gamma}(\boldsymbol{\Lambda}\mathbf{S}_k)^{1/2} \right]' \boldsymbol{\Gamma}(\boldsymbol{\Lambda}\mathbf{S}_k)^{1/2} \mathbf{S}_k \\ &= (\boldsymbol{\Lambda}\mathbf{S}_k)^{1/2} \boldsymbol{\Gamma}' \boldsymbol{\Gamma} (\boldsymbol{\Lambda}\mathbf{S}_k)^{1/2} \mathbf{S}_k \\ &= \boldsymbol{\Lambda}\mathbf{S}_k\mathbf{S}_k = \boldsymbol{\Lambda},\end{aligned}$$

puesto que los vectores propios en $\boldsymbol{\Gamma}$ eran ortonormales. Así, $\mathbf{X}'\mathbf{X}\mathbf{S}_k$ es sencillamente la matriz diagonal de valores propios no nulos de la matriz \mathbf{B} , trivialmente invertible. En este caso la Ecuación (6.9) para el cálculo del vector \mathbf{x} queda simplificada:

$$\mathbf{x} = \frac{1}{2} \boldsymbol{\Lambda}^{-1} \mathbf{X}' \mathbf{d}. \quad (6.10)$$

Sólo falta determinar la última componente, $\mathbf{x}_{P(d+1)}$, del nuevo vector, que se puede obtener de la Ecuación (6.8):

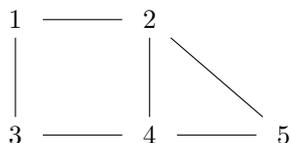
$$\mathbf{x}_{P(d+1)} = \pm \sqrt{\left| \mathbf{x}' \mathbf{S}_k \mathbf{x} + \frac{1}{n} \mathbf{1}' \mathbf{d} \right|}. \quad (6.11)$$

Notar que si esta cantidad es nula y \mathbf{X} era no degenerada, entonces el nuevo conjunto de puntos degenera y no es necesario añadir la nueva dimensión. Además, se puede tomar tanto positiva como negativa.

En resumen, dada la matriz \mathbf{X} de coordenadas originales no degenerada y el vector $\boldsymbol{\delta}$ de distancias adicionales, si \mathbf{X} está en coordenadas principales se obtiene \mathbf{x} con la Fórmula (6.10), en caso contrario se utiliza la Fórmula (6.9). De la Fórmula (6.8) se tiene que el signo asociado a la nueva dimensión es el signo de la cantidad $-(\mathbf{x}' \mathbf{S}_k \mathbf{x} + \frac{1}{n} \mathbf{1}' \mathbf{d})$ y la coordenada del nuevo punto en la nueva dimensión viene dada por (6.11).

6.2.3 Ejemplos

Consideraremos la siguiente configuración esquemática en la cual las distancias se miden sólo a lo largo de las aristas.



Los dos primeros puntos están separados por una distancia igual a 1. Considerados aisladamente, tienen una matriz de distancias y una representación Euclídea en un espacio de dimensión 1 dadas respectivamente por

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \mathbf{x} = \begin{bmatrix} -0,5 \\ 0,5 \end{bmatrix}.$$

Si a esta configuración se añade el tercer punto, a distancias $\delta = (1, 2)$ de los dos primeros puntos respectivamente, la matriz de distancias y una representación que mantiene las coordenadas de los puntos originales son, respectivamente,

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{pmatrix}; \quad \mathbf{x}^+ = \begin{bmatrix} -0,5 \\ 0,5 \\ -1,5 \end{bmatrix}.$$

Notar que esta es una representación degenerada, ya que no ha sido necesario incrementar la dimensionalidad de la representación. La representación continúa teniendo lugar en la recta Euclídea. Los puntos 1 y 2 preservan las coordenadas obtenidas en la representación anterior. En cambio, si se calcula la representación de los tres puntos conjuntamente se obtiene la representación centrada

$$\mathbf{x} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}.$$

En este punto, no es posible añadir un punto más a esta representación porque es degenerada, y las fórmulas que hemos calculado parten de la base de una configuración no-degenerada. Con las herramientas que hemos desarrollado sólo es posible calcular la representación de los cuatro puntos conjuntamente.

Supongamos que se añade el cuarto punto, con distancias $\boldsymbol{\delta} = (2, 1, 1)$ a los anteriores. Ahora esta configuración de distancias deja de ser Euclídea. Se trata precisamente del ejemplo de [Curriero \(2006\)](#) que introdujimos en la Sección 3.2.3. Sin embargo existe una representación pseudo-Euclídea con índice $k = 2$, dada a continuación junto con la matriz de distancias:

$$\begin{pmatrix} 0 & 1 & 1 & 2 \\ 1 & 0 & 2 & 1 \\ 1 & 2 & 0 & 1 \\ 2 & 1 & 1 & 0 \end{pmatrix}; \quad \mathbf{X} = \begin{bmatrix} 0 & 1 & \overset{(-)}{0,5} \\ 1 & 0 & -0,5 \\ -1 & 0 & -0,5 \\ 0 & -1 & 0,5 \end{bmatrix}.$$

Notar que la dimensionalidad ha aumentado en dos unidades al añadir un único punto, debido a que la anterior representación era degenerada. Ahora la dimensionalidad se ha ubicado en su valor *natural* $d = n - 1$.

Se añade ahora el quinto punto, con distancias $\boldsymbol{\delta} = (1 + \sqrt{2}, \sqrt{2}, 2, 1)$ a los anteriores. La representación pseudo-Euclídea (de índice $k = 3$) que se obtiene conservando las coordenadas de la representación anterior, y la representación centrada que se obtiene a partir de todas las distancias conjuntamente, son las que siguen

$$\mathbf{X}^+ = \begin{bmatrix} 0 & 1 & 0 & \overset{(-)}{0,5} \\ 1 & 0 & 0 & -0,5 \\ -1 & 0 & 0 & -0,5 \\ 0 & -1 & 0 & 0,5 \\ 0,5 & -1,21 & 0,89 & 0,21 \end{bmatrix}; \quad \mathbf{X} = \begin{bmatrix} -1,21 & 0,38 & 0,15 & \overset{(-)}{0,51} \\ 0,09 & 0,92 & -0,18 & -0,53 \\ -0,64 & -0,92 & 0,05 & -0,51 \\ 0,56 & -0,38 & -0,39 & 0,45 \\ 1,20 & 0 & 0,37 & 0,08 \end{bmatrix}.$$

Notar que en la representación que conserva las coordenadas anteriores se ha producido un reordenamiento de las columnas, ya que la nueva dimensión se ha incorporado con signo positivo.

6.2.4 Relación con la Relatividad Especial

En la Teoría de la Relatividad Especial de Einstein, tanto las distancias como los intervalos de tiempo dependen de la velocidad del observador. Este postulado permitió explicar un fenómeno para el que no se tenía respuesta: el hecho de que la velocidad de las ondas electromagnéticas fuera constante independientemente de la velocidad del observador, algo que contradice el postulado de *Relatividad Galileana* y la concepción absoluta del espacio y el tiempo de Newton.

Einstein desarrolló algebraicamente las consecuencias de este postulado en su famoso artículo de 1905. Dos años más tarde, el matemático Hermann Minkowski, que fuera profesor de matemáticas de Einstein, mostró que la Relatividad Especial podía entenderse geoméricamente si se consideraba el espacio-tiempo como un espacio de 4 dimensiones en el cual se midieran *intervalos espacio-temporales* de acuerdo con la ecuación

$$ds^2 = dx_1^2 + dx_2^2 + dx_3^2 - dt^2,$$

donde dx_i representan distancias espaciales en cada una de las tres dimensiones, y dt representa un intervalo temporal. De esta manera, independientemente de su velocidad relativa, diferentes observadores medirían el mismo intervalo espacio-temporal ds .

El espacio-tiempo de Minkowski es, en efecto, un espacio pseudo-Euclídeo $E_{3,1}$. Las analogías resultan fascinantes. El intervalo espacio-temporal es precisamente la distancia pseudo-Euclídea.

En el espacio de Minkowski, los puntos de cuadrado positivo se dice que tienen una separación tipo-espacio con respecto al origen. Esto significa que no se puede viajar (a una velocidad menor que la de la luz) desde el origen hasta

dichos puntos. También significa que aunque el punto en cuestión se encuentre en el *futuro*, un observador viajando a la velocidad adecuada percibiría dichos puntos como sucesos *simultáneos*.

En el espacio de Minkowski las funciones isotrópicas se llaman *invariantes-Lorentz*, o invariantes ante transformaciones de Lorentz. Las transformaciones de Lorentz son las *rotaciones* del espacio pseudo-Euclídeo, es decir, transformaciones que conservan la distancia pseudo-Euclídea. Los puntos de un hiperboloide *rotan* de modo que siempre quedan sobre la misma superficie, así como los puntos de una esfera rotan sobre la misma esfera en el espacio Euclídeo. En particular, los puntos de cuadrado positivo no pueden pasar a la región de puntos de cuadrado negativo mediante una transformación de Lorentz. De otra manera, no se hubiera conservado la distancia al origen.

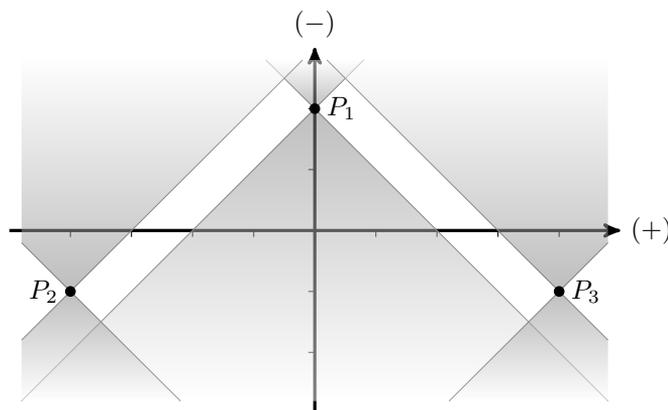
6.3 Definición Positiva en los espacios pseudo-Euclídeos

En la sección anterior hemos establecido la posibilidad de representar cualquier problema de distancias basadas en el coste en un espacio pseudo-Euclídeo. Esto puede verse como una inmersión del espacio métrico basado en el coste D (ver Sec. 5.1) en el espacio pseudo-Euclídeo.

Es importante notar que esta inmersión no es estricta. Es decir, no se trata de una identificación. El espacio pseudo-Euclídeo es más amplio, tiene configuraciones de puntos que no son la representación de ningún problema de distancias basadas en el coste.

Esto es una consecuencia directa de que el espacio pseudo-Euclídeo no es un espacio *métrico*, sino *pre-métrico*. Específicamente, la premétrica pseudo-Euclídea no cumple las propiedades de *identidad de indiscernibles* ($d(\mathbf{x}, \mathbf{y}) = 0 \iff \mathbf{x} = \mathbf{y}$) ni la *desigualdad triangular* ($d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$)

Consideremos por ejemplo la siguiente configuración de tres puntos en el espacio $E_{1,1}$.



En las unidades del gráfico, la distancia pseudo-Euclídea entre el punto P_1 y cualquiera de los otros dos es $\vartheta(P_1, P_2) = \vartheta(P_1, P_3) = \sqrt{4^2 - 3^2} = \sqrt{7}$, mientras que entre P_2 y P_3 la distancia pseudo-Euclídea es igual a 8, coincidiendo con la distancia Euclídea, ya que ambos puntos se encuentran sobre el mismo eje de signo positivo. Esta configuración viola claramente la desigualdad triangular, ya que $8 = \vartheta(P_2, P_3) \not\leq \vartheta(P_2, P_1) + \vartheta(P_1, P_3) = 2\sqrt{7} \approx 5,3$.

Se concluye que esta configuración, aunque está centrada, no puede ser la representación de ninguna estructura de distancias basadas en el coste. Con este contraejemplo se demuestra que, aunque todo problema de distancias basadas en el coste se puede representar en el espacio pseudo-Euclídeo, lo recíproco no es cierto.

Desde el punto de vista práctico, necesitamos ahora encontrar familias de funciones válidas (i.e., definidas positivas) en este espacio para poder utilizarlas como funciones de correlación en aplicaciones reales.

Ya sabemos que no todas las funciones válidas en el espacio Euclídeo lo son también en el pseudo-Euclídeo. Vimos en la Sección 3.2.3, por ejemplo, que la función de correlación Gaussiana no era válida en general.

Puesto que el espacio pseudo-Euclídeo es un espacio más general, que contiene al Euclídeo como caso particular (a saber, cuando $k = d$), necesariamente la

familia de funciones válidas deben ser un subconjunto de aquellas que son válidas en el espacio Euclídeo.

Sabemos también que este subconjunto es no-vacío, ya que existen funciones válidas en contextos incluso más generales. Un ejemplo es la función trivial constante, mencionada en la Sección 5.4. La función exponencial es una candidata que plausiblemente pueda pertenecer a este conjunto, ya que ha producido matrices de correlación definidas positivas tanto en el ejemplo de Curriero (Sec. 3.2.3) como en el proyecto piloto (Sec. 4.8).

Esquemáticamente, podemos decir que el espacio Euclídeo E es una particularización del espacio métrico basado en el coste D , y éste a su vez se puede representar en un espacio pseudo-Euclídeo $E_{(k,d-k)}$, que es aún más general.

Consecuentemente, las respectivas familias de funciones definidas positivas de cada espacio mantienen las relaciones inversas de inclusión. Abusando de la notación:

$$\emptyset \neq \mathcal{P}(E_{(k,d-k)}) \subseteq \mathcal{P}(D) \subseteq \mathcal{P}(E).$$

Entendiendo que las funciones sobre D se aplican en realidad sobre la representación de sus puntos en el espacio pseudo-Euclídeo.

Mientras que en el capítulo anterior nos concentramos en la familia $\mathcal{P}(D)$, que es realmente la que nos interesa, en este capítulo intentamos caracterizar la sub-familia $\mathcal{P}(E_{(k,d-k)})$. Aunque es sólo una parte de lo que nos interesa, la estructura casi Euclidiana del espacio puede facilitar el trabajo.

El objetivo más ambicioso es encontrar caracterizaciones análogas a las de Bochner y Schoenberg (ver Sec. 2.3), que proporcionen (todas las) funciones de correlación definidas positivas en problemas de distancias basadas en el coste.

En segundo lugar investigaremos en particular la función exponencial, para intentar determinar si es definida positiva en todo el espacio pseudo-Euclídeo y si pertenece, por tanto, al conjunto de funciones válidas.

Empezaremos por notar que la representación espectral del Teorema de Bochner (Teo. 2.4, p.21) es aplicable al espacio pseudo-Euclídeo, en tanto que

se refiere a funciones definidas en \mathbb{R}^d . Si bien es cierto que en la representación de Bochner y en su demostración se utiliza el producto y la norma Euclídea, no hay nada que impida aplicar estas funciones en el espacio pseudo-Euclídeo, con la única salvedad de que no coinciden con el pseudo-producto interno y la pseudo-métrica definidos en el espacio.

De modo que, efectivamente, conocemos todas las funciones definidas positivas del espacio pseudo-Euclídeo: son exactamente las mismas que las del espacio Euclídeo. Esto tiene además la interesante consecuencia de que, por las relaciones de inclusión discutidas más arriba, este mismo conjunto de funciones es el conjunto de funciones válidas para el espacio métrico D , representado como espacio pseudo-Euclídeo.

La diferencia se manifiesta en la subclase de funciones *isotrópicas*, es decir, aquellas que sólo dependen de la *distancia* entre localizaciones. En efecto, las funciones isotrópicas en el espacio Euclídeo son constantes sobre las superficies esféricas, ya que la esfera es el lugar geométrico de los puntos que distan del origen una cierta cantidad. Por el contrario, en el espacio pseudo-Euclídeo, las superficies de distancia constante al origen son hiperboloides. En la Figura 6.3 se representan estas superficies para los espacios de dimensión 3 E_3 y $E_{2,1}$.

Así pues, aunque las funciones definidas positivas sean las mismas en ambos espacios, las que son isotrópicas en el espacio pseudo-Euclídeo no son isotrópicas en el espacio Euclídeo y viceversa. La única excepción la constituye, trivialmente, la función constante en todo el espacio, que lógicamente es isotrópica bajo cualquier punto de vista.

Recuérdese de la Sección 2.3 que para encontrar la forma de las funciones isotrópicas en el espacio Euclídeo, Schoenberg integraba la representación de Bochner sobre la superficie de una esfera. Así, aprovechaba el hecho de que la función era constante allí y la mayor parte del trabajo consistía en calcular la integral de la función exponencial $e^{i\omega'x}$ sobre la superficie esférica, con respecto al vector ω para un x fijo. Esto se hacía factible en coordenadas hipersféricas, de modo que el producto $\omega'x$ sólo dependía del radio y del primero de los ángulos hipersféricos, lo que permitía integrar el resto de los ángulos independientemente.

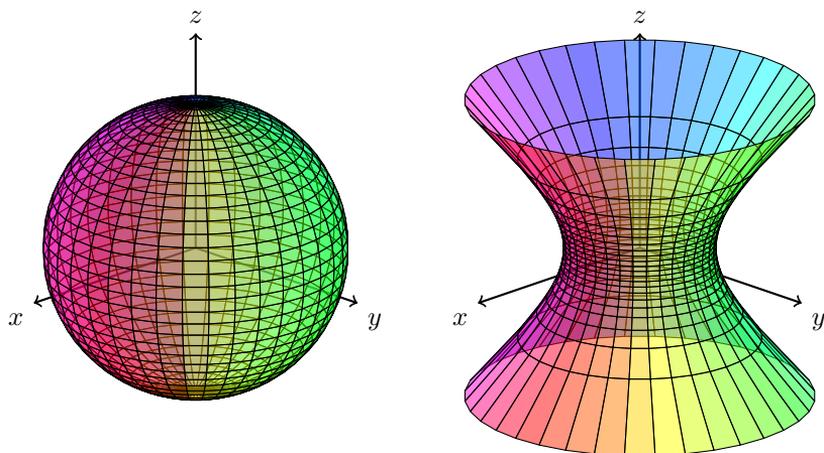


Figura 6.3. Superficies isotrópicas del espacio Euclídeo y pseudo-Euclídeo

En el espacio pseudo-Euclídeo seguiremos los mismos pasos, pero ahora será necesario integrar sobre el d -hiperboloide. Para lo cual necesitamos un sistema de coordenadas adaptado a la geometría pseudo-Euclídea.

A estos efectos, en el Apéndice F desarrollamos un sistema de coordenadas que llamamos *pseudo-hiperesféricas*, en el entendido de que el hiperboloide se puede considerar como la esfera de los espacios pseudo-Euclídeos.

Realmente, en el espacio pseudo-Euclídeo existen dos esferas de radio ρ . La formada por puntos de cuadrado positivo ρ^2 y la formada por puntos de cuadrado negativo $-\rho^2$. En ambos casos, la distancia al origen de los puntos en dichas superficies es ρ , de acuerdo con la definición de distancia pseudo-Euclídea (6.4).

Las parametrizaciones de las regiones de cuadrado positivo y negativo son diferentes. En ambos casos, este sistema cuenta con un parámetro ρ que representa el *radio* del hiperboloide, es decir, la distancia pseudo-Euclídea constante de cada uno de los puntos de la superficie al origen de coordenadas.

Para la región de cuadrado positivo, se cuenta además con $k - 1$ ángulos esféricos ϕ_2, \dots, ϕ_k . El primero, ϕ_2 es un ángulo azimutal que varía en $[0, 2\pi)$ y

que proporciona la simetría de revolución. Los demás varían entre $-\pi/2$ y $\pi/2$. Por último, hay $d - k$ ángulos hiperbólicos $\alpha_{k+1}, \dots, \alpha_d$ variando en todo \mathbb{R} .

El jacobiano de la transformación de coordenadas pseudo-hiperesféricas a cartesianas $T_\rho^{(+)}(\phi_2, \dots, \phi_k, \alpha_{k+1}, \dots, \alpha_d)$, en esta región, viene dado por

$$|\det J(T_\rho^{(+)})| = |\rho^{d-1} \prod_{j=3}^k (\cos \phi_j)^{j-2} \prod_{j=k+1}^d (\cosh \alpha_j)^{j-2}|.$$

En cambio, en la región de cuadrado negativo, se tienen k ángulos hiperbólicos $\alpha_1, \dots, \alpha_k$ y $d - k - 1$ ángulos esféricos $\phi_{k+1}, \dots, \phi_{d-1}$, el último de los cuales es azimutal y varía en $[0, 2\pi)$.

El jacobiano de la correspondiente transformación de puntos en esta región es

$$|\det J(T_\rho^{(-)})| = |\rho^{d-1} \prod_{j=1}^k (\cosh \alpha_j)^{d-j-1} \prod_{j=k+1}^{d-2} (\cos \phi_j)^{d-j-1}|.$$

Hay una consideración más que vale la pena mencionar aquí. Estas parametrizaciones son válidas para espacios pseudo-Euclídeos $E_{(k,d-k)}$ tales que $2 \leq k < k + 2 \leq d$. Los casos extremos $k = 1$ y $d - k = 1$ requieren una parametrización alternativa, que también produce sendos jacobianos específicos. Referimos al Apéndice F sobre éste y otros detalles del sistema de coordenadas pseudo-hiperesférico.

Surge naturalmente la pregunta de si es necesario definir las funciones isotrópicas como aquellas que son constantes sobre *ambas superficies* a la vez, considerándolas como dos componentes desconectadas de una única esfera.

No tenemos una respuesta definitiva, ya que el problema de las funciones definidas positivas en espacios pseudo-Euclídeos sigue abierto y nos acercamos a la frontera de lo que hemos conseguido por esta línea de trabajo. Sin embargo creemos que es suficiente considerar funciones constantes sobre la esfera de cuadrado positivo. El argumento que sostiene este criterio es que el Teorema 6.6 de representación pseudo-Euclídea asegura que los puntos que se encuentran siempre subtienden distancias cuadráticas positivas (ver comentario al final

de la demostración). Por tanto, a los efectos del cálculo geoestadístico basado en el coste, las funciones constantes sobre la esfera de cuadrado positivo son efectivamente isotrópicas.

Quisiéramos reproducir a continuación el argumento del Teorema 2.5 de Schoenberg (p. 25), en el contexto del espacio pseudo-Euclídeo. Esencialmente se reduce a integrar sobre la superficie esférica de radio h la representación espectral de Bochner de una función de correlación isotrópica

$$R(\mathbf{x}) = \int_{\mathbb{R}^d} e^{i\boldsymbol{\omega}'\mathbf{x}} F(d\boldsymbol{\omega}),$$

donde F es una medida positiva finita. Como la función de correlación R es isotrópica, el lado izquierdo es simplemente $R(h)$ por la superficie de la esfera. El lado derecho resulta una integral sobre todo el espacio, con respecto a la medida finita F , de un integrando que sólo depende de $\|\boldsymbol{\omega}\|$, por lo que se puede poner como una integral unidimensional sobre el semieje positivo con respecto a una nueva medida finita G definida a partir de F .

A continuación ofrecemos resultados parciales, ya que no hemos podido completar el cálculo. En cualquier caso, queremos dejar constancia de los pequeños avances que hemos podido hacer en esta dirección y establecer explícitamente las dificultades que han quedado sin resolver.

En el espacio pseudo-Euclídeo, las superficies isotrópicas tienen superficie infinita (son hiperboloides), por lo que integrar la representación de Bochner sobre ellas conduce a una divergencia. Podemos, sin embargo, considerar el valor medio de la función R sobre la superficie isotrópica de radio ρ , de cuadrado positivo

$$S_\rho^+ = \{\mathbf{x} \in \mathbb{R}^d : x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_d^2 = \rho^2\}.$$

Puesto que R es constante sobre dicha superficie, el valor medio es igual al valor en cualquiera de sus puntos y sólo depende del radio ρ . El cálculo del valor medio del lado derecho sólo se puede expresar formalmente como el cociente de

dos integrales divergentes:

$$\int_{S_\rho^+} \left(\int_{\mathbb{R}^d} e^{i\boldsymbol{\omega}'\mathbf{x}} F(d\boldsymbol{\omega}) \right) s(d\mathbf{x}) \Big/ \int_{S_\rho^+} s(d\mathbf{x}) = \\ \int_{\mathbb{R}^d} \left(\int_{S_\rho^+} e^{i\boldsymbol{\omega}'\mathbf{x}} s(d\mathbf{x}) \Big/ \int_{S_\rho^+} s(d\mathbf{x}) \right) F(d\boldsymbol{\omega}).$$

Cambiando el orden de integración resulta que $R(\rho)$ se puede calcular como la integral respecto de la medida F de la función definida formalmente como

$$M_{\|\boldsymbol{\omega}\|} \{e^{i\boldsymbol{\omega}'\mathbf{x}}\}(\rho) = \int_{S_\rho^+} e^{i\boldsymbol{\omega}'\mathbf{x}} s(d\mathbf{x}) \Big/ \int_{S_\rho^+} s(d\mathbf{x}).$$

Esto es, el valor medio de la función exponencial sobre la esfera de radio ρ . Naturalmente depende también de $\boldsymbol{\omega}$ que se considera fijo, pero sólo a través de su módulo, como argumentaremos enseguida.

Si consiguiéramos una expresión analítica para esta función bastaría reescribir la integral en términos de una medida G sobre $[0, \infty)$, igual que en el Teorema de Schoenberg, y obtendríamos así una representación de las funciones definidas positivas isotrópicas en el espacio pseudo-Euclídeo.

6.3.1 Resultados parciales sobre la función $M(\rho)$

En primer lugar nos preguntamos si tiene sentido calcular el valor medio de la función $e^{i\boldsymbol{\omega}'\mathbf{x}}$ sobre una superficie infinita. Si la función creciera indefinidamente, por ejemplo, simplemente no existiría un valor medio. Sin embargo, notemos que esta función es acotada. Concretamente, tiene módulo 1. Esto no demuestra que exista un valor medio, pero al menos no lo descarta.

Cuando el dominio de integración es infinito, como en este caso, la integral impropia se define a través de un proceso de límite. Es común que el valor resulte diferente según cómo se haga el proceso de límite. Un método para asignar un

valor unívocamente consiste en utilizar el Valor Principal de Cauchy (CPV, por sus siglas en inglés).

Para el cálculo de la función $M(\rho)$, necesitamos fijar un vector $\boldsymbol{\omega}$, que puede ser de cuadrado positivo o negativo (o cero, pero este conjunto de casos tiene medida nula). Si lo suponemos de cuadrado positivo, podemos elegir sin pérdida de generalidad el vector con componentes $\omega_i = \delta_{ik}\|\boldsymbol{\omega}\|$, donde δ es la función delta de Kronecker. Esto es, un vector en la dirección del eje k . De otro modo, siempre se puede hacer una rotación (hiperbólica) que lleve un vector $\boldsymbol{\omega}$ de cuadrado positivo a la dirección especificada (ver Sec. F.4). Por tanto, se tiene

$$\boldsymbol{\omega}'\boldsymbol{x} = \|\boldsymbol{\omega}\|x_k = \|\boldsymbol{\omega}\| \sin \phi_k \prod_{j=3}^d \cosh \alpha_j,$$

escribiendo x_k en las coordenadas pseudo-hiperesféricas correspondientes a la región de cuadrado positivo.

Así, la función exponencial no depende de los primeros $k - 2$ ángulos esféricos y basta hacer la media sobre el resto de las dimensiones. Formalmente,

$$\int_{S_\rho^+} e^{i\boldsymbol{\omega}'\boldsymbol{x}} s(d\boldsymbol{x}) = \int_{\Phi} \exp\{i\|\boldsymbol{\omega}\| \sin \phi_k \prod_{j=k+1}^d \cosh \alpha_j\} \rho^{d-1} \prod_{j=3}^d |c_j|^{j-2} d\phi_2 \cdots d\phi_k d\alpha_{k+1} \cdots d\alpha_d,$$

donde $\Phi = [0, 2\pi) \times [-\frac{\pi}{2}, \frac{\pi}{2}] \times \mathbb{R}^{d-k}$ y c_j es la notación definida en (F.2). De aquí se puede extraer un factor

$$\rho^{d-1} \int \prod_{j=3}^{k-1} |\cos \phi_j|^{j-2} d\phi_2 \cdots d\phi_{k-1}$$

que se cancela con el correspondiente en $\int_{S_\rho^+} s(d\boldsymbol{x})$.

Lo que, llamando $A = \|\boldsymbol{\omega}\| \prod_{j=k+1}^d \cosh \alpha_j$, reduce la función M a

$$M_{\|\boldsymbol{\omega}\|} \{e^{i\boldsymbol{\omega}'\mathbf{x}}\}(\rho) = \frac{\int e^{iA \sin \phi_k} (\cos \phi_k)^{k-2} \prod_{j=k+1}^d (\cosh \alpha_j)^{j-2} d\phi_k d\alpha_{k+1} \cdots d\alpha_d}{\int (\cos \phi_k)^{k-2} \prod_{j=k+1}^d (\cosh \alpha_j)^{j-2} d\phi_k d\alpha_{k+1} \cdots d\alpha_d}$$

Aún podemos integrar el último ángulo esférico. En efecto,

$$\int_{[-\frac{\pi}{2}, \frac{\pi}{2}]} e^{iA \sin \phi_k} (\cos \phi_k)^{k-2} d\phi_k = \int_{[-\frac{\pi}{2}, \frac{\pi}{2}]} \cos(A \sin \phi_k) (\cos \phi_k)^{k-2} d\phi_k + i \int_{[-\frac{\pi}{2}, \frac{\pi}{2}]} \sin(A \sin \phi_k) (\cos \phi_k)^{k-2} d\phi_k.$$

La componente imaginaria es cero, pues el integrando es una función impar en ϕ_k . Por su parte, el integrando de la componente real es par y la integral puede hacerse en el intervalo $[0, \pi/2]$. Además, si $k \geq 3$, con el cambio de variable

$$x = \sin \phi_k,$$

se tiene $dx = \cos \phi_k d\phi_k$ y $\cos \phi_k = \sqrt{1-x^2}$, por tanto,

$$\begin{aligned} \int_{[-\frac{\pi}{2}, \frac{\pi}{2}]} e^{iA \sin \phi_k} (\cos \phi_k)^{k-2} d\phi_k &= 2 \int_0^{\frac{\pi}{2}} (\cos \phi_k)^{k-2} \cos(A \sin \phi_k) d\phi_k \\ &= 2 \int_0^1 (1-x^2)^{\frac{k-3}{2}} \cos(Ax) dx \\ &= \sqrt{\pi} \left(\frac{2}{A}\right)^{\frac{k-2}{2}} \Gamma\left(\frac{k-1}{2}\right) J_{\frac{k-2}{2}}(A), \end{aligned}$$

donde J_ν es la función de Bessel de primera especie (Abramowitz y Stegun 1972, Ec. 9.1.20, Gradshteyn y Ryzhik 2007, Sec. 8.4).

En este último paso hemos utilizado el siguiente resultado que se encuentra en [Gradshteyn y Ryzhik \(2007, 3.771\(8\)\)](#):

$$\int_0^u (u^2 - x^2)^{\nu - \frac{1}{2}} \cos(ax) dx = \frac{\sqrt{\pi}}{2} \left(\frac{2u}{a}\right)^\nu \Gamma\left(\nu + \frac{1}{2}\right) J_\nu(au),$$

con $a > 0$, $u > 0$, $|\Re \nu| > -\frac{1}{2}$. En nuestro caso las condiciones se verifican pues $a = A = \|\omega\|\rho \prod_{j=k+1}^d \cosh \alpha_j > 0$; $u = 1 > 0$ y $\nu = \frac{k-2}{2} > -\frac{1}{2}$, pues $k \geq 2$.

Por su parte, ϕ_k también puede integrarse en el denominador de la función M , pues

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (\cos \phi_k)^{k-2} d\phi_k = 2 \int_0^{\frac{\pi}{2}} (\cos \phi_k)^{k-2} d\phi_k = \frac{\sqrt{\pi} \Gamma\left(\frac{k-1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)}.$$

De modo que, cancelando los factores comunes $\sqrt{\pi}$ y $\Gamma\left(\frac{k-1}{2}\right)$ en el numerador y el denominador, se tiene la siguiente expresión de la función M :

$$M_{\|\omega\|} \{e^{i\omega'x}\}(\rho) = \left(\frac{2}{\|\omega\|\rho}\right)^{\frac{k-2}{2}} \Gamma\left(\frac{k}{2}\right) \frac{\int_{\mathbb{R}^{d-k}} J_{\frac{k-2}{2}}(A) \prod_{j=k+1}^d (\cosh \alpha_j)^{j - \frac{k}{2} - 1} d\alpha_{k+1} \cdots d\alpha_d}{\int_{\mathbb{R}^{d-k}} \prod_{j=k+1}^d (\cosh \alpha_j)^{j-2} d\alpha_{k+1} \cdots d\alpha_d}$$

con $A = \|\omega\|\rho \prod_{j=k+1}^d \cosh \alpha_j$. También puede notarse que como las variables α_j siempre aparecen como argumentos de Cosenos hiperbólicos, ambos integrandos son funciones pares con respecto a cada variable, por lo que la integración podría hacerse sobre los intervalos $[0, \infty)$ introduciendo un factor 2. Esto resulta en sendos factores 2^{d-k} en el numerador y en el denominador que se cancelan exactamente.

Aquí es donde se presenta el problema de la divergencia, ya que hay que integrar los ángulos hiperbólicos sobre un dominio infinito. Naturalmente esta

expresión es sólo formal. Nuestra esperanza era conseguir *cancelar* factores divergentes y resolver así la indeterminación.

El problema se puede reducir algo más notando que el numerador se puede resolver de forma anidada, una variable a la vez. Por su parte, el denominador se puede poner como producto de las integrales de cada uno de los factores.

De esta manera, se puede trabajar con una serie de integrales unidimensionales anidadas, la primera de las cuales sería, por ejemplo, para α_{k+1} :

$$\frac{\int_0^\infty J_{\frac{k-2}{2}} \left(\|\omega\| \rho \prod_{j=k+1}^d \cosh \alpha_j \right) (\cosh \alpha_{k+1})^{\frac{k}{2}} d\alpha_{k+1}}{\int_0^\infty (\cosh \alpha_{k+1})^{k-1} d\alpha_{k+1}}.$$

El resultado se integraría a su vez respecto de α_{k+2} y así sucesivamente, en una sucesión de operaciones que tienen la misma estructura.

Haciendo el cambio de variable $x = \cosh \alpha_{k+1}$ y denotando $A_s = \|\omega\| \rho \prod_{k+s}^d$ se tiene

$$\frac{\int_1^\infty x^{\frac{k}{2}} J_{\frac{k}{2}-1}(A_2 x) \frac{dx}{\sqrt{x^2-1}}}{\int_1^\infty x^{k-1} \frac{dx}{\sqrt{x^2-1}}}.$$

La Figura 6.4 muestra el aspecto del integrando del numerador para dos casos particulares de k . Aunque la función tiene un módulo claramente creciente (tanto

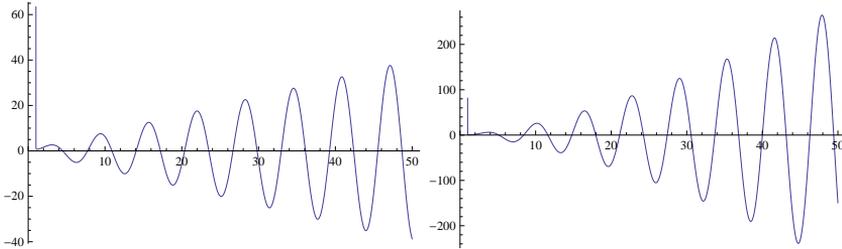


Figura 6.4. Curvas $\frac{x^{\frac{k}{2}}}{\sqrt{x^2-1}} J_{\frac{k}{2}-1}(Ax)$ para $k = 5$ (izq.) y $k = 6$ (der.).

más cuando mayor es k), el programa de cálculo simbólico *Mathematica* ([Wolfram Research, Inc., 2008](#)) resuelve la integral como

$$\frac{\sqrt{\pi/2}}{\sqrt{A_2}} \sec\left(\frac{k\pi}{2}\right) \left[J_{\frac{k-1}{2}}(A_2) \sin\left(\frac{k\pi}{2}\right) - J_{\frac{1-k}{2}}(A_2) \right].$$

Notar que la función Secante se hace infinita toda vez que k es impar. Y de hecho, resultan valores finitos cuando se computa el valor numéricamente.

Sin embargo, cuando se incluye este resultado en el siguiente nivel, para integrar α_{k+2} ya nos encontramos con una integral que no podemos resolver:

$$\int_0^\infty J_{\frac{k-1}{2}}(A_3 \cosh \alpha_{k+2}) (\cosh \alpha_{k+2})^{\frac{k+1}{2}} d\alpha_{k+2} = \int_1^\infty x^{\frac{k+1}{2}} J_{\frac{k-1}{2}}(A_3 x) \frac{dx}{\sqrt{x^2-1}}.$$

Realmente esperábamos desde el principio que estas integrales fueran divergentes. Nuestra esperanza residía en resolver el cociente conjuntamente, como el límite de las integrales propias,

$$\lim_{L \rightarrow \infty} \frac{\int_1^L x^{\frac{k}{2}} J_{\frac{k}{2}-1}(A_2 x) \frac{dx}{\sqrt{x^2-1}}}{\int_1^L x^{k-1} \frac{dx}{\sqrt{x^2-1}}}$$

de modo que se cancelaran las componentes divergentes.

De hecho, *Mathematica* es capaz de calcular el denominador como

$$\frac{\sqrt{\pi} \Gamma\left(\frac{1-k}{2}\right)}{2\Gamma\left(\frac{2-k}{2}\right)} + \frac{L^{k-1} {}_2F_1\left[\frac{1}{2}, \frac{1-k}{2}, \frac{3-k}{2}, \frac{1}{L^2}\right]}{k-1},$$

donde ${}_2F_1$ es la función Hipergeométrica de Gauss ([Abramowitz y Stegun, 1972](#), Sec. 15). Sin embargo, no hemos sido capaces de resolver la integral en el numerador dividido por esta cantidad.

Por si fuera de utilidad para el futuro, nos gustaría señalar que en el esfuerzo por resolver estas integrales también hemos intentado reformularlas sobre un intervalo de longitud finita, cosa que se puede hacer con el cambio de variable

$$y = \operatorname{arcsec} x,$$

con lo cual, $dy = \frac{dx}{x\sqrt{x^2-1}}$. De esta manera, el intervalo de integración pasa a ser $[0, \pi/2]$ aunque lógicamente el integrando tiende a infinito cuando y se acerca a $\pi/2$. Con este cambio, sin embargo, no hemos obtenido mayores avances.

Por último, también hemos considerado la posibilidad de utilizar una representación de la función de Bessel como una serie de Potencias ([Abramowitz y Stegun, 1972](#), Ec. 9.1.10):

$$\begin{aligned} J_\nu(z) &= (z/2)^\nu \sum_{s=0}^{\infty} \frac{(-1)^s z^{2s}}{4^s \Gamma(\nu + s + 1) s!} \\ &= \frac{(z/2)^\nu}{\Gamma(\nu + 1)} \left[1 - \frac{z^2}{4(\nu + 1)} + \frac{z^4}{32(\nu + 1)(\nu + 2)} - \dots \right]. \end{aligned}$$

Con lo cual,

$$\begin{aligned} \int_1^{\infty} x^{\frac{k}{2}} J_{\frac{k}{2}-1}(Ax) \frac{dx}{\sqrt{x^2-1}} &= \int_1^{\infty} \frac{x^{\frac{k}{2}}}{\sqrt{x^2-1}} \frac{(Ax/2)^{\frac{k}{2}-1}}{\Gamma(\frac{k}{2})} \left[1 - \frac{(Ax)^2}{2k} + \dots \right] dx \\ &= \frac{(A/2)^{\frac{k}{2}-1}}{\Gamma(\frac{k}{2})} \int_1^{\infty} \frac{x^{k-1}}{\sqrt{x^2-1}} \left[1 - \frac{(Ax)^2}{2k} + \dots \right] dx. \end{aligned}$$

Y dividiendo por el denominador se tendría una serie de cocientes de integrales:

$$\frac{(A/2)^{\frac{k}{2}-1}}{\Gamma(\frac{k}{2})} \left[1 + \frac{\int_1^{\infty} \frac{(Ax)^2}{2k}}{\int_1^{\infty} \frac{dx}{x^{k-1} \sqrt{x^2-1}}} + \dots \right].$$

De todos modos, el problema del cociente de integrales divergentes persiste. Aunque ahora está notablemente simplificado, tenemos un número infinito de ellos que hay que resolver en general y luego escribir la serie resultante en términos de alguna función conocida, posiblemente otra función de Bessel. Luego aún quedaría pasar al siguiente ángulo hiperbólico α_{k+2} y repetir el proceso.

Con idénticos resultados hemos probado otras estrategias como cambiar el orden de integración, integrando en primer lugar las variables hiperbólicas, utilizar la representación integral de la función de Bessel o utilizar alguna de sus propiedades de recurrencia.

Como se ve, hay un problema de fondo que es el del cálculo del valor medio de una función sobre un dominio infinito. No importa cómo se escriba nuestro problema, o cuánto se simplifique y reduzca. Al final siempre persiste esta dificultad.

Hemos dejado planteados todos los elementos con los que hemos trabajado en esta línea. En el futuro, si es que encontramos un método para resolver este tipo de cálculos, podemos retomar aquella posibilidad que resulte más conveniente. O quizás alguien más, con mayor pericia en el Análisis, algún día lea estas líneas y pueda completar estos cálculos que proporcionarían una importantísima representación de una familia de funciones definidas positivas e isotrópicas en el espacio pseudo-Euclídeo.

6.4 Cálculo de funciones de densidad espectral

Ante las dificultades en la búsqueda de una representación general para todas las funciones definidas positivas isotrópicas en el espacio pseudo-Euclídeo, nos planteamos el objetivo más modesto de encontrar alguna función particular y de estudiar la definición positiva de algunas funciones de correlación estándar.

6.4.1 Densidad espectral de la función de correlación constante en $E_{(k,d-k)}$

Hemos afirmado que este conjunto de funciones definidas positivas e isotrópicas en el espacio pseudo-Euclídeo, que denotamos por $\mathcal{P}(E_{(k,d-k)})$ es no vacío. En efecto, hemos demostrado con una simple comprobación que la función constante igual a 1 es definida positiva y ciertamente es isotrópica.

Según el Teorema de Bochner, esta función tendrá una representación espectral. Es decir, existe una medida finita F tal que

$$1 \equiv R(\mathbf{x}) = \int_{\mathbb{R}^d} e^{i\boldsymbol{\omega}'\mathbf{x}} F(d\boldsymbol{\omega}).$$

En efecto, basta considerar la medida que asigna toda su *masa* al punto $\boldsymbol{\omega} = \mathbf{0}$, pues entonces la integral se reduce a la evaluación del integrando en ese punto y $e^{i\mathbf{0}'\mathbf{x}} = 1$. Esta medida es conocida como la δ de Dirac. Se trata de un caso límite, para el cual no existe la función de densidad espectral.

En el espacio Euclídeo, éste es un caso límite de la función de correlación exponencial cuando el parámetro φ tiende a cero. En efecto,

$$\lim_{\varphi \rightarrow 0} e^{-\varphi h} = 1, \quad \forall h = \|\mathbf{x}\| \geq 0.$$

Por su parte, su función de densidad espectral (Ec. 2.11)

$$f(\boldsymbol{\omega}) = \frac{\Gamma(\frac{d+1}{2})}{\pi^{\frac{d+1}{2}}} \frac{\varphi}{(\varphi^2 + \|\boldsymbol{\omega}\|^2)^{\frac{d+1}{2}}}, \quad d \geq 1$$

tiene límite

$$\lim_{\varphi \rightarrow 0} f(\boldsymbol{\omega}) = \begin{cases} 0 & \|\boldsymbol{\omega}\| > 0 \\ \infty & \|\boldsymbol{\omega}\| = 0. \end{cases}$$

Es decir, precisamente la *función* δ de Dirac, que sería la función de densidad espectral correspondiente a la *medida* δ si nos tomamos la licencia de asignar el

valor infinito al cero, situándonos en el terreno de las funciones generalizadas, o *distribuciones*.

Puesto que el caso límite es una función definida positiva también en el espacio pseudo-Euclídeo, resulta natural preguntarse si la función exponencial es también una función válida allí. Por otra parte, las pruebas empíricas que hemos hecho en los experimentos aplicados no han descartado la definición positiva de esta función, al menos en el espacio métrico basado en el coste.

De forma preliminar, deduciremos las fórmulas para el cálculo de la densidad espectral de una función de correlación isotrópica genérica. Para simplificar y concretar los cálculos, trabajaremos específicamente en el $E_{(2,1)}$, que es precisamente el espacio donde hemos representado el ejemplo de Curriero (Sec. 6.2.3). Posteriormente aplicaremos estas fórmulas al caso exponencial.

6.4.2 Densidad espectral de una función de correlación isotrópica en $E_{(2,1)}$

Una función de correlación isotrópica en $E_{(2,1)}$ por definición depende únicamente de la distancia pseudo-Euclídea al origen de coordenadas. Es decir, es una función $R(\mathbf{x})$ tal que (abusando de la notación)

$$R(\mathbf{x}) = R(x_1^2 + x_2^2 - x_3^2) = R(\rho^2),$$

donde ρ es el parámetro radial en el sistema de coordenadas pseudo-hiperesféricas (ver Ap. F).

La función de densidad espectral se calcula como la transformada inversa de Fourier

$$f(\boldsymbol{\omega}) = \frac{1}{(2\pi)^3} \iiint_{\mathbb{R}^3} R(x_1^2 + x_2^2 - x_3^2) e^{-i\boldsymbol{\omega}'\mathbf{x}} dx_1 dx_2 dx_3.$$

Igual que hicimos en el planteamiento de la generalización de Schoenberg, en la sección anterior, fijado $\boldsymbol{\omega}$ podemos rotar el espacio pseudo-Euclídeo para que

este vector sea colineal con uno de los ejes coordenados. Por la naturaleza de las rotaciones en el espacio pseudo-Euclídeo tenemos que distinguir los casos en que $\boldsymbol{\omega}$ sea de cuadrado positivo o negativo. Así, si $(\boldsymbol{\omega}, \boldsymbol{\omega}) = \omega_1^2 + \omega_2^2 - \omega_3^2 > 0$, podemos asumir sin pérdida de generalidad que $\omega_1 = \omega_3 = 0$ y $\omega_2 = \|\boldsymbol{\omega}\| > 0$ (ver Sec. F.4).

La integración tiene que extenderse sobre todo el espacio, pero en coordenadas pseudo-hiperesféricas necesitamos dos parametrizaciones diferentes, una para la región de cuadrado positivo y otra para la región de cuadrado negativo (ver Ap. F). Particularizadas para el caso $E_{(2,1)}$, las parametrizaciones son las siguientes:

$$\begin{cases} x_1 = \rho \cos \phi \cosh \alpha \\ x_2 = \rho \sin \phi \cosh \alpha \\ x_3 = \rho \sinh \alpha \end{cases} \quad \begin{cases} x_1 = \rho \sinh \alpha \cos \phi \\ x_2 = \rho \sinh \alpha \sin \phi \\ x_3 = \rho \cosh \alpha \end{cases}$$

para las regiones

$$I = \{\boldsymbol{x} : (\boldsymbol{x}, \boldsymbol{x}) > 0\} = \{(\rho, \phi, \alpha) \in [0, \infty) \times [0, 2\pi] \times \mathbb{R}\}$$

y

$$II = \{\boldsymbol{x} : (\boldsymbol{x}, \boldsymbol{x}) < 0\} = \{(\rho, \phi, \alpha) \in \mathbb{R} \times [-\frac{\pi}{2}, \frac{\pi}{2}] \times \mathbb{R}\},$$

con jacobianos $\rho^2 \cosh \alpha$ y $\rho^2 |\sinh \alpha|$ respectivamente. Para quitar el valor absoluto en el último, podemos integrar α en $[0, -\infty) \cup [0, \infty)$, ya que precisamente cuando $\alpha < 0$, se tiene $\sinh \alpha < 0$.

Con todo esto, se tiene:

$$f(\boldsymbol{\omega}) = \frac{1}{(2\pi)^3} \left[\int_I R(\rho^2) e^{-i\|\boldsymbol{\omega}\| \rho \sin \phi \cosh \alpha} \rho^2 \cosh \alpha \, d\rho d\phi d\alpha + \int_{II} R(\rho^2) e^{-i\|\boldsymbol{\omega}\| \rho \sin \phi \sinh \alpha} \rho^2 |\sinh \alpha| \, d\rho d\phi d\alpha \right].$$

• **Región I.**

Podemos integrar el ángulo ϕ en primer lugar. Anotando $A = \|\boldsymbol{\omega}\|\rho \cosh \alpha$ y haciendo el cambio de variable $\phi = 2\pi - \tilde{\phi}$ en la segunda mitad del intervalo de integración,

$$\begin{aligned} \int_0^{2\pi} e^{-iA \sin \phi} d\phi &= \int_0^{\pi} e^{-iA \sin \phi} d\phi + \int_{\pi}^{2\pi} e^{-iA \sin \tilde{\phi}} d\tilde{\phi} \\ &= \int_0^{\pi} e^{-iA \sin \phi} + e^{+iA \sin \phi} d\phi \\ &= 2 \int_0^{\pi} \cos(A \sin \phi) d\phi. \end{aligned}$$

Repitiendo el truco, dividimos el intervalo en dos y hacemos el cambio $\phi = \pi - \tilde{\phi}$, con lo cual $\sin \phi = \sin \tilde{\phi}$ y $\int_{\pi/2}^{\pi} \cos(A \sin \tilde{\phi}) d\tilde{\phi} = -\int_{\pi/2}^0 \cos(A \sin \phi) d\phi$. Por tanto,

$$\int_0^{2\pi} e^{-iA \sin \phi} d\phi = 4 \int_0^{\pi/2} \cos(A \sin \phi) d\phi.$$

Finalmente, utilizando el cambio de variable $x = \sin \phi$, y la integral 3.753(2) del [Gradshteyn y Ryzhik \(2007\)](#),

$$\int_0^{2\pi} e^{-i\|\boldsymbol{\omega}\|\rho \cosh \alpha \sin \phi} d\phi = 2\pi J_0(\|\boldsymbol{\omega}\|\rho \cosh \alpha),$$

donde J_0 es la función de Bessel de primera especie de orden 0.

Así, la integral sobre la región I queda, teniendo en cuenta que $\cosh \alpha$ es par,

$$4\pi \int_0^{\infty} R(\rho^2) \rho^2 \int_0^{\infty} \cosh \alpha J_0(\|\boldsymbol{\omega}\|\rho \cosh \alpha) d\alpha d\rho$$

La integración de α se puede hacer con el cambio de variable $x = \cosh \alpha$, de modo que (Gradshteyn y Ryzhik, 2007, Ec. 6.554(3)),

$$\begin{aligned} \int_0^\infty \cosh \alpha J_0(\|\boldsymbol{\omega}\|\rho \cosh \alpha) d\alpha &= \int_1^\infty x J_0(\|\boldsymbol{\omega}\|\rho x) \frac{dx}{\sqrt{x^2 - 1}} \\ &= \frac{\cos(\|\boldsymbol{\omega}\|\rho)}{\|\boldsymbol{\omega}\|\rho}. \end{aligned}$$

Con lo cual sólo nos queda la siguiente integral en ρ :

$$\frac{4\pi}{\|\boldsymbol{\omega}\|} \int_0^\infty R(\rho^2) \rho \cos(\|\boldsymbol{\omega}\|\rho) d\rho.$$

• **Región II.**

Ajustemos primero los dominios. La integral en α puede hacerse en el semieje positivo haciendo el cambio $\tilde{\alpha} = -\alpha$ en el semieje negativo:

$$\int_{\mathbb{R}} |\sinh \alpha| e^{-iA \sinh \alpha} d\alpha = \int_0^\infty \sinh \alpha \underbrace{(e^{-iA \sinh \alpha} + e^{+iA \sinh \alpha})}_{=2 \cos(A \sinh \alpha)} d\alpha.$$

Además, el integrando es par en ϕ y en ρ . Por lo tanto,

$$\begin{aligned} \int_{-\infty}^\infty R(\rho^2) \rho^2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left(\int_{-\infty}^\infty |\sinh \alpha| e^{-i\|\boldsymbol{\omega}\|\rho \sin \phi \sinh \alpha} d\alpha \right) d\phi d\rho = \\ 8 \int_0^\infty R(\rho^2) \rho^2 \int_0^\infty \sinh \alpha \int_0^{\frac{\pi}{2}} \cos(\|\boldsymbol{\omega}\|\rho \sin \phi \sinh \alpha) d\phi d\alpha d\rho. \end{aligned}$$

La integral en ϕ se resuelve utilizando el cambio $x = \sin \phi$ y la integral 3.753(2) del Gradshteyn y Ryzhik (2007), que ya apareció en la integración de la región I, resultando en $\pi/2 J_0(A)$. Con lo que se tiene:

$$4\pi \int_0^\infty R(\rho^2) \rho^2 \int_0^\infty \sinh \alpha J_0(\|\boldsymbol{\omega}\|\rho \sinh \alpha) d\alpha d\rho.$$

Ahora, con el cambio $x = \sinh \alpha$ y la integral 6.554(1) del [Gradshteyn y Ryzhik \(2007\)](#), Con lo cual sólo nos queda la siguiente integral en ρ :

$$\frac{4\pi}{\|\boldsymbol{\omega}\|} \int_0^\infty R(\rho^2) \rho e^{-\|\boldsymbol{\omega}\| \rho} d\rho.$$

En resumen, en el espacio $E_{(2,1)}$ la densidad espectral de una función isotrópica se puede calcular, para $\|\boldsymbol{\omega}\| = \sqrt{|(\boldsymbol{\omega}, \boldsymbol{\omega})|}$ y $(\boldsymbol{\omega}, \boldsymbol{\omega}) > 0$, como

$$f(\boldsymbol{\omega}) = \frac{1}{2\pi^2 \|\boldsymbol{\omega}\|} \int_0^\infty R(\rho^2) \rho \left(\cos(\|\boldsymbol{\omega}\| \rho) + e^{-\|\boldsymbol{\omega}\| \rho} \right) d\rho.$$

Calculamos ahora la expresión de la densidad espectral para $\boldsymbol{\omega}$ de cuadrado negativo. Se puede asumir, sin pérdida de generalidad, $\omega_1 = \omega_2 = 0$; $\omega_3 = \|\boldsymbol{\omega}\| > 0$. Basta para ello aplicar una rotación adecuada del sistema de coordenadas y eventualmente una inversión de la tercera dimensión (ver Sec. [F.4](#)).

Nuevamente tenemos que integrar separadamente en las regiones I y II , para calcular la expresión de la densidad espectral para $(\boldsymbol{\omega}, \boldsymbol{\omega}) < 0$:

$$f(\boldsymbol{\omega}) = \frac{1}{(2\pi)^3} \left[\int_I R(\rho^2) \rho^2 \cosh \alpha e^{-i\|\boldsymbol{\omega}\| \rho \sinh \alpha} d\rho d\phi d\alpha + \int_{II} R(\rho^2) \rho^2 |\sinh \alpha| e^{-i\|\boldsymbol{\omega}\| \rho \cosh \alpha} d\rho d\phi d\alpha \right].$$

• **Región I.**

Integrando trivialmente el ángulo ϕ se tiene

$$2\pi \int_0^\infty R(\rho^2) \rho^2 \int_{-\infty}^\infty \cosh \alpha e^{-i\|\boldsymbol{\omega}\| \rho \sinh \alpha} d\alpha d\rho.$$

La integral interior en α no converge. Notar que cambiando el signo de la variable de integración,

$$\int_{-\infty}^0 \cosh \alpha e^{-iA \sinh \alpha} d\alpha = \int_0^\infty \cosh \alpha e^{iA \sinh \alpha} d\alpha,$$

lo que sumado a la otra mitad de la integral y haciendo el cambio $x = \sinh \alpha$, resulta

$$\begin{aligned} \int_0^\infty \cosh \alpha \left(e^{iA \sinh \alpha} + e^{-iA \sinh \alpha} \right) d\alpha &= 2 \int_0^\infty \cosh \alpha \cos(\|\boldsymbol{\omega}\| \rho \sinh \alpha) d\alpha \\ &= 2 \int_0^\infty \cos(\|\boldsymbol{\omega}\| \rho x) dx, \end{aligned}$$

cuyo integrando es oscilante y la integral no converge.

Sin embargo, notar que $\forall \varepsilon \approx 0$,

$$\begin{aligned} \int_{-\infty}^\infty \cosh(\alpha - i\varepsilon) e^{-iA \sinh(\alpha - i\varepsilon)} d\alpha &= -\frac{1}{iA} \left[e^{-iA \sinh(\alpha - i\varepsilon)} \right]_{-\infty}^\infty \\ &= -\frac{1}{iA} \left[\lim_{\alpha \rightarrow \infty} \underbrace{e^{-iA \sinh \alpha}}_{\text{acotado}} \underbrace{e^{-\varepsilon A \cosh \alpha}}_{\rightarrow 0} - \lim_{\alpha \rightarrow -\infty} \underbrace{e^{-iA \sinh \alpha}}_{\text{acotado}} \underbrace{e^{-\varepsilon A \cosh \alpha}}_{\rightarrow 0} \right] = 0. \end{aligned}$$

Aquí hemos utilizado la identidad $\sinh(a \pm ib) = \sinh a \cos b \pm i \cosh a \sin b$, de modo que si $b = \varepsilon \rightarrow 0$, entonces $\cos \varepsilon \rightarrow 1$ mientras que $\sin \varepsilon \rightarrow 0$, por lo que $\sinh(\alpha - i\varepsilon) \approx \sinh \alpha - i\varepsilon \cosh \alpha$.

De modo que se puede asignar a la integral el valor 0, por ser el límite de integrales que tienden a 0 (Wurm et al., 2003).

• Región II.

Integrando trivialmente el ángulo ϕ y utilizando la paridad en α se tiene

$$2\pi \int_{-\infty}^\infty R(\rho^2) \rho^2 \int_0^\infty \sinh \alpha e^{-i\|\boldsymbol{\omega}\| \rho \cosh \alpha} d\alpha d\rho.$$

Aquí nuevamente ocurre que la integral en α no es convergente, pero la podemos plantear como (Wurm et al., 2003)

$$\lim_{\varepsilon \rightarrow 0} \int_0^\infty \sinh(\alpha - i\varepsilon) e^{-iA \cosh \alpha - i\varepsilon} d\alpha = \lim_{\varepsilon \rightarrow 0} -\frac{1}{iA} \left[e^{-iA \cosh(\alpha - i\varepsilon)} \right]_{\alpha=0}^{\alpha \rightarrow \infty}.$$

Ahora, teniendo en cuenta que $\cosh(a \pm ib) = \cosh a \cos b \pm i \sinh a \sin b$, cuando $\varepsilon \approx 0$ se tiene $\cosh(\alpha - i\varepsilon) \approx \cosh \alpha - i\varepsilon \sinh \alpha$. De modo que el límite vale

$$-\frac{1}{iA} \lim_{\varepsilon \rightarrow 0} \left[\lim_{\alpha \rightarrow \infty} \underbrace{e^{-iA \cosh \alpha}}_{\text{acotado}} \underbrace{e^{-\varepsilon A \sinh \alpha}}_{\rightarrow 0} - e^{-iA \cosh(i\varepsilon)} \right] = \frac{e^{-iA}}{iA}$$

siempre que $A = \|\boldsymbol{\omega}\| \rho > 0$, es decir, cuando $\rho > 0$. En caso contrario, hay que considerar $\alpha + i\varepsilon$, lo que proporciona el mismo límite. Por tanto, sustituyendo en la integral global se tiene

$$\frac{2\pi}{i\|\boldsymbol{\omega}\|} \int_{-\infty}^{\infty} R(\rho^2) \rho e^{-i\|\boldsymbol{\omega}\| \rho} d\rho.$$

Finalmente, haciendo el cambio de variable $\tilde{\rho} = -\rho$ en el semieje negativo, resulta

$$\frac{2\pi}{i\|\boldsymbol{\omega}\|} \int_0^{\infty} R(\rho^2) \rho \underbrace{\left(e^{-i\|\boldsymbol{\omega}\| \rho} - e^{i\|\boldsymbol{\omega}\| \rho} \right)}_{-2i \sin(\|\boldsymbol{\omega}\| \rho)} d\rho = \frac{-4\pi}{\|\boldsymbol{\omega}\|} \int_0^{\infty} R(\rho^2) \rho \sin(\|\boldsymbol{\omega}\| \rho) d\rho.$$

Puesto que la integración sobre la región I es nula, se tiene la siguiente expresión de la densidad espectral cuando $\boldsymbol{\omega}$ es de cuadrado negativo:

$$f(\boldsymbol{\omega}) = \frac{-1}{2\pi^2 \|\boldsymbol{\omega}\|} \int_0^{\infty} R(\rho^2) \rho \sin(\|\boldsymbol{\omega}\| \rho) d\rho.$$

En resumen, en el espacio $E_{(2,1)}$ la densidad espectral de una función isotrópica se puede calcular, con las siguientes fórmulas:

$$f(\boldsymbol{\omega}) = \begin{cases} \frac{1}{2\pi^2 \|\boldsymbol{\omega}\|} \int_0^{\infty} R(\rho^2) \rho \left(\cos(\|\boldsymbol{\omega}\| \rho) + e^{-\|\boldsymbol{\omega}\| \rho} \right) d\rho, & (\boldsymbol{\omega}, \boldsymbol{\omega}) > 0 \\ \frac{-1}{2\pi^2 \|\boldsymbol{\omega}\|} \int_0^{\infty} R(\rho^2) \rho \sin(\|\boldsymbol{\omega}\| \rho) d\rho, & (\boldsymbol{\omega}, \boldsymbol{\omega}) < 0 \end{cases} \quad (6.12)$$

Curiosamente, estas integrales se corresponden con la suma de las transformadas de Laplace y la transformada Coseno de Fourier en el primer caso y con la transformada Seno de Fourier en el segundo.

Notar que hemos partido de una función de correlación isotrópica, aplicado el Teorema de Bochner y llegado a una expresión que sólo depende de la premétrica pseudo-Euclídea (i.e., del parámetro ρ). Esto es general, y es una consecuencia de la posibilidad de rotación del sistema de coordenadas. En efecto, si no importa la orientación inicial del vector ω (dentro de la región correspondiente), se deriva inmediatamente que el resultado sólo puede depender de su magnitud. En otras palabras, si la función de correlación es isotrópica, la densidad espectral (si existe) también lo es.

6.4.3 Densidad espectral de la función de correlación exponencial en $E_{(2,1)}$

En esta sección nos proponemos calcular la función de densidad espectral correspondiente a la función de correlación exponencial en el espacio pseudo-Euclídeo. Por el Teorema de Bochner, si (y sólo si) la función resultante es no-negativa entonces la función exponencial será una función definida positiva.

La función de correlación exponencial tiene la misma forma que en el caso Euclídeo, excepto por el hecho de que la norma del vector \mathbf{x} se sustituye por la premétrica pseudo-Euclídea, para reflejar la dependencia con la distancia pseudo-Euclídea. Vista como una función multivariante, la función exponencial en $E_{(2,1)}$ es entonces

$$e^{-\varphi\sqrt{|x_1^2+x_2^2-x_3^2|}}, \quad \varphi > 0.$$

Para realizar el cálculo es conveniente trabajar en coordenadas pseudo-hiperesféricas (ver Ap. F), en cuyos términos la función adopta la expresión

$$e^{-\varphi\sqrt{\rho^2}} = e^{-\varphi|\rho|},$$

y podemos utilizar las fórmulas deducidas en la sección anterior.

La Transformada Coseno resulta:

$$\int_0^{\infty} \rho e^{-\varphi\rho} \cos(\|\boldsymbol{\omega}\|\rho) d\rho = \frac{1}{\|\boldsymbol{\omega}\|^2 + \varphi^2} \cos(2 \arctan(\|\boldsymbol{\omega}\|/\varphi)).$$

según [Gradshteyn y Ryzhik \(2007, Ec. 3.944\(6\)\)](#).

Se puede escribir más elegantemente, utilizando el Coseno del ángulo doble y escribiendo los cuadrados del Seno y el Coseno en términos de la tangente: $\cos^2 x = 1/(1 + \tan^2 x)$; $\sin^2 x = \tan^2 x/(1 + \tan^2 x)$. De ese modo,

$$\begin{aligned} \cos(2 \arctan A) &= \cos^2(\arctan A) - \sin^2(\arctan A) \\ &= \frac{1}{1 + A^2} - \frac{A^2}{1 + A^2}. \end{aligned}$$

Quedando entonces:

$$\int_0^{\infty} \rho e^{-\varphi\rho} \cos(\|\boldsymbol{\omega}\|\rho) d\rho = \frac{\varphi^2 - \|\boldsymbol{\omega}\|^2}{(\varphi^2 + \|\boldsymbol{\omega}\|^2)^2}.$$

La Transformada de Laplace es:

$$\int_0^{\infty} \rho e^{-(\varphi + \|\boldsymbol{\omega}\|)\rho} d\rho = \frac{1}{(\varphi + \|\boldsymbol{\omega}\|)^2}.$$

Finalmente, necesitamos calcular la Transformada Seno:

$$\int_0^{\infty} \rho e^{-\varphi\rho} \sin(\|\boldsymbol{\omega}\|\rho) d\rho = \frac{1}{\|\boldsymbol{\omega}\|^2 + \varphi^2} \sin(2 \arctan(\|\boldsymbol{\omega}\|/\varphi)) = \frac{2\varphi\|\boldsymbol{\omega}\|}{(\varphi^2 + \|\boldsymbol{\omega}\|^2)^2}$$

según [Gradshteyn y Ryzhik \(2007, Ec. 3.944\(5\)\)](#) y reescribiendo la expresión con el Seno del ángulo doble en términos de la función Tangente.

En resumen, en el espacio $E_{(2,1)}$ la densidad espectral de la función de correlación exponencial es:

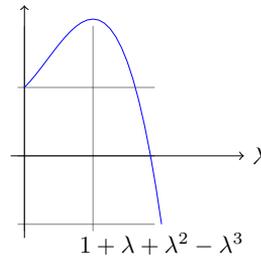
$$f(\boldsymbol{\omega}) = \begin{cases} \frac{1}{2\pi^2\|\boldsymbol{\omega}\|} \left(\frac{\varphi^2 - \|\boldsymbol{\omega}\|^2}{(\varphi^2 + \|\boldsymbol{\omega}\|^2)^2} + \frac{1}{(\varphi + \|\boldsymbol{\omega}\|)^2} \right), & (\boldsymbol{\omega}, \boldsymbol{\omega}) > 0 \\ \frac{-\varphi}{\pi^2(\varphi^2 + \|\boldsymbol{\omega}\|^2)^2}, & (\boldsymbol{\omega}, \boldsymbol{\omega}) < 0 \end{cases}$$

con $\|\boldsymbol{\omega}\| = \sqrt{|(\boldsymbol{\omega}, \boldsymbol{\omega})|}$.

Notar que esta función no es no-negativa en toda la región de cuadrado positivo. Esto implica que la función exponencial *no es definida positiva* en el $E_{(2,1)}$.

En efecto, si medimos la magnitud de la frecuencia en unidades de φ , $\|\boldsymbol{\omega}\| = \lambda\varphi$, entonces la función, para $(\boldsymbol{\omega}, \boldsymbol{\omega}) > 0$ adopta la siguiente forma:

$$f(\lambda) = \frac{1}{\pi^2\varphi^3} \frac{1 + \lambda + \lambda^2 - \lambda^3}{\lambda(1 + \lambda)^2(1 + \lambda^2)^2}$$



Con lo cual, eventualmente se hace negativa para $\|\boldsymbol{\omega}\|$ suficientemente grande (de hecho, apenas un par de veces la magnitud del parámetro φ). Aunque cuando $\lambda \rightarrow \infty$, la función se acerca a 0^- .

6.4.4 Expresiones alternativas de la función de correlación exponencial en $E_{(2,1)}$

Una posible estrategia para evitar el problema de indefinición positiva es modificar el comportamiento de la función de correlación exponencial en la región II. Aprovechando el hecho (Teo. 6.6) de que las representaciones pseudo-Euclídeas

de problemas basados en el coste siempre subtienden distancias de cuadrado positivo, un cambio de definición en esta región no afectaría a las correlaciones que intervienen en el problema. De esa manera, eligiendo convenientemente la expresión de la función en esa región podría eventualmente conseguirse que se cancelara la componente que se hace negativa en la expresión de la densidad espectral correspondiente a la integración en la región I .

Un ejemplo de esto, que no es solución, sería definir la función de correlación, para $\varphi > 0$, como

$$C(\mathbf{x}) = e^{-\varphi\sqrt{x_1^2+x_2^2-x_3^2}} = \begin{cases} e^{-\varphi\rho}, & x_1^2 + x_2^2 - x_3^2 = \rho^2 > 0 \\ e^{-i\varphi|\rho|}, & x_1^2 + x_2^2 - x_3^2 = -\rho^2 < 0. \end{cases}$$

De este modo, la integración sobre la región I (la transformada Coseno) es igual, porque la función es la misma. Por su parte la integración sobre la región II , que corresponde a la transformación de Laplace, resulta:

$$\int_0^\infty \rho e^{-(\|\boldsymbol{\omega}\|+i\varphi)\rho} d\rho = \frac{1}{(\|\boldsymbol{\omega}\|+i\varphi)^2} = \left(\frac{\|\boldsymbol{\omega}\| - i\varphi}{\|\boldsymbol{\omega}\|^2 + \varphi^2} \right)^2.$$

Uniendo las dos componentes, resulta, para $(\boldsymbol{\omega}, \boldsymbol{\omega}) > 0$:

$$f(\boldsymbol{\omega}) = \frac{-i\varphi}{\pi^2(\varphi^2 + \|\boldsymbol{\omega}\|^2)^2},$$

lo cual, naturalmente, no resuelve el problema.

Sirva como ejemplo del tipo de manipulación por el cual, modificando la definición de la función de correlación en la región de cuadrado negativo se puede alterar la forma de la función de densidad espectral.

Otra definición con un resultado interesante es $C(\mathbf{x}) = \cos(\varphi\rho)$, en la región II . Produce una situación totalmente simétrica entre $\|\boldsymbol{\omega}\|$ y φ , al tener que sumar la transformada Coseno de la función exponencial y la transformada de Laplace de la función Coseno. De modo que la suma en las dos regiones se cancela

exactamente y la función de densidad espectral, para $(\boldsymbol{\omega}, \boldsymbol{\omega}) > 0$, es nula en todas partes. Como se ve, no es necesario restringir la definición de la función a la *forma* exponencial.

Esto sugiere inmediatamente definir la función exponencial como

$$C(\boldsymbol{x}) = \begin{cases} e^{-\varphi\rho}, & x_1^2 + x_2^2 - x_3^2 = \rho^2 > 0 \\ \cos(\varphi\rho) + e^{-\varphi|\rho|}, & x_1^2 + x_2^2 - x_3^2 = -\rho^2 < 0. \end{cases}$$

De manera que la transformada de Laplace resulta

$$\int_0^\infty \rho \left(\cos(\varphi\rho) + e^{-\varphi\rho} \right) e^{-\|\boldsymbol{\omega}\|\rho} d\rho = \left[\underbrace{\int_0^\infty \rho \cos(\varphi\rho) e^{-\|\boldsymbol{\omega}\|\rho} d\rho}_{=\frac{\|\boldsymbol{\omega}\|^2 - \varphi^2}{(\|\boldsymbol{\omega}\|^2 + \varphi^2)^2}} + \underbrace{\int_0^\infty \rho e^{-(\varphi + \|\boldsymbol{\omega}\|)\rho} d\rho}_{=\frac{1}{(\varphi + \|\boldsymbol{\omega}\|)^2}} \right].$$

El primer término cancela entonces el que resulta de la transformada Coseno, con lo cual, resulta finalmente

$$f(\boldsymbol{\omega}) = \frac{1}{2\pi^2 \|\boldsymbol{\omega}\|} \frac{1}{(\varphi + \|\boldsymbol{\omega}\|)^2} > 0, \quad (\boldsymbol{\omega}, \boldsymbol{\omega}) > 0,$$

con $\|\boldsymbol{\omega}\| = \sqrt{|(\boldsymbol{\omega}, \boldsymbol{\omega})|}$.

Para el caso $(\boldsymbol{\omega}, \boldsymbol{\omega}) < 0$ tenemos que aplicar la transformada Seno a la función tal como está definida en la región *II*, multiplicada por ρ . Esto es,

$$\int_0^\infty \rho \left(\cos(\varphi\rho) + e^{-\varphi\rho} \right) \sin(\|\boldsymbol{\omega}\|\rho) d\rho = \int_0^\infty \rho \cos(\varphi\rho) \sin(\|\boldsymbol{\omega}\|\rho) d\rho + \int_0^\infty \rho e^{-\varphi\rho} \sin(\|\boldsymbol{\omega}\|\rho) d\rho.$$

El primer término no es integrable. El producto de Seno y Coseno se puede poner en términos de los Senos de la suma y diferencia de los argumentos, con

lo que la integral se desdobra en dos:

$$\frac{1}{2} \left[\int_0^\infty \rho \sin((\|\boldsymbol{\omega}\| + \phi)\rho) d\rho + \int_0^\infty \rho \sin((\|\boldsymbol{\omega}\| - \phi)\rho) d\rho \right]$$

que son claramente oscilantes y divergentes. La única asignación arbitraria razonable para esa integral es cero, ya que cualquier valor resultante de integrar hasta cierto punto se cancela eventualmente.

Para el segundo término utilizamos la integral 3.944(5) del [Gradshteyn y Ryzhik \(2007\)](#):

$$\int_0^\infty \rho e^{-\varphi\rho} \sin(\|\boldsymbol{\omega}\|\rho) d\rho = \frac{\sin(2 \arctan \frac{\|\boldsymbol{\omega}\|}{\varphi})}{\|\boldsymbol{\omega}\|^2 + \varphi^2}.$$

Teniendo en cuenta que $\sin(2 \arctan x) = 2x/(1+x^2)$ se tiene finalmente,

$$f(\boldsymbol{\omega}) = \frac{-\varphi}{\pi^2(\|\boldsymbol{\omega}\|^2 + \varphi^2)^2}$$

lo cual es claramente negativo y por tanto descarta que la función de correlación exponencial, así definida, sea definida positiva.

En conclusión, hemos presentado en esta sección y en la anterior nuestros esfuerzos por comprobar si la función de correlación exponencial es definida positiva en el $E_{(2,1)}$. Hemos demostrado que no lo es, al menos en su expresión más *natural*. También hemos observado que tenemos cierta *libertad* de cambiar su expresión en la región de cuadrado negativo, sin cambiar la estructura de correlaciones de un problema basado en el coste. Esto nos permite intentar ajustar esta expresión para intentar una densidad espectral no-negativa. Hemos conseguido *arreglar* la densidad espectral en la región de cuadrado positivo, pero no en la de cuadrado negativo. A pesar de nuestros esfuerzos, no hemos encontrado una definición de la cual resulte una función de densidad espectral no-negativa simultáneamente en ambas regiones del espacio. Por lo que, de

momento, debemos aceptar que la función de correlación exponencial, tal cual la hemos definido, no es definida positiva en el $E_{(2,1)}$.

Notar que esto tampoco constituye una prueba de que la función exponencial no sea válida para problemas basados en el coste.

6.5 Conclusiones

Comenzamos este capítulo con una introducción al Multidimensional Scaling (MDS) como una posible aproximación al trabajo con distancias basadas en el coste. Esta técnica permite obtener una representación Euclídea de la estructura de distancias. Sin embargo, esto se consigue a costa de filtrar las particularidades de la región, que son precisamente la motivación de las distancias basadas en el coste.

Aunque no nos satisface como solución, el MDS nos sirvió de inspiración para el desarrollo de la representación pseudo-Euclídea. Esta representación permite una representación *exacta* (Teo. 6.6) de una estructura de distancias no necesariamente Euclídea, en un espacio vectorial dotado de un *producto interno indefinido*. Esto es, una forma bilineal simétrica pero en la que la condición de definición positiva es sustituida por la de no-degeneración (Def. 6.5).

Este producto induce una *pre métrica simétrica* en el espacio pseudo-Euclídeo que, entre otras cosas, no verifica la desigualdad triangular. Sin embargo la distancia basada en el coste sí cumple con esta propiedad. Esto significa que la representación pseudo-Euclídea de todos los posibles problemas basados en el coste no agotan todas las configuraciones posibles de puntos en el espacio pseudo-Euclídeo.

Por ejemplo, ninguna representación de un problema basado en el coste tendrá dos puntos alineados sobre uno de los ejes de signo negativo. Eso significaría que existirían dos puntos en el problema original con una distancia cuadrática negativa, lo que es imposible. Otra configuración imposible es la analizada en el ejemplo al inicio de la Sección 6.3. Incluso siendo todas las distancias positivas, se viola la desigualdad triangular.

En este sentido, el espacio pseudo-Euclídeo es más *grande* que el espacio métrico D ; aunque es sabido que, tratándose de conjuntos infinitos, la inclusión de uno dentro de otro no es un criterio adecuado para comparar su tamaño. Esto parecería indicar que la propiedad de definición positiva es más exigente en el espacio pseudo-Euclídeo, que en el métrico D . Una función aspirante a definida positiva en el espacio pseudo-Euclídeo tiene que cumplir la propiedad en cuestión para una gran variedad de configuraciones adicionales, con configuraciones de puntos que ni siquiera tienen por qué cumplir la desigualdad triangular.

Sin embargo, hemos argumentado que el conjunto $\mathcal{P}(D)$ de funciones multivariantes definidas positivas en el espacio D debe ser el mismo que en el espacio pseudo-Euclídeo (y que en el Euclídeo).

La complejidad sobreviene cuando nos ocupamos de la subfamilia de funciones isotrópicas, que es la que verdaderamente nos interesa desde el punto de vista de las aplicaciones. La obtención de la representación espectral de una función isotrópica exige calcular el valor medio de una función exponencial compleja sobre una superficie hiperboloide, de extensión infinita.

Para poder integrar en el espacio pseudo-Euclídeo fue necesario desarrollar un sistema de coordenadas (Ap. F) adaptado al producto pseudo-Euclídeo. Es imprescindible considerar por separado las regiones de puntos de cuadrado positivo y negativo en el espacio pseudo-Euclídeo, ya que la parametrización es diferente en cada una.

El cálculo del valor medio de una función sólo puede definirse con seguridad en un dominio de medida finita. En caso contrario, el resultado puede ser diferente dependiendo de cómo se lleve a cabo el proceso de límite. Es habitual utilizar el *Valor Principal de Cauchy* como criterio para asignar valores a integrales impropias de este tipo. Aún así, sólo hemos sido capaces de obtener resultados parciales en este cálculo. Más precisamente, hemos sido capaces de integrar los ángulos esféricos que varían en un dominio de medida finita y de reducir el problema de integración impropia múltiple a una única dimensión. Pero finalmente no conseguimos resolver la integral, aunque hemos esbozado las estrategias de resolución ensayadas.

Puesto que no pudimos avanzar más en el análisis general de las funciones definidas positivas isotrópicas, nos planteamos el estudio de casos particulares de funciones y sus densidades espectrales respectivas.

Comenzamos analizando la función de correlación constante, que sabemos que es definida positiva y, por supuesto, isotrópica. Vimos que es un caso límite para el cual no existe función de densidad espectral, cuya representación espectral se basa en la medida δ de Dirac.

Posteriormente calculamos las fórmulas de la densidad espectral de una función isotrópica en el espacio $E_{(2,1)}$, que nos sirvieron de base para el cálculo de la densidad espectral de la función exponencial en el mismo espacio. Llegamos a la inesperada conclusión de que la función exponencial *no es* definida positiva en $E_{(2,1)}$ y por tanto no está legitimada, en principio, como función de correlación para un problema basado en el coste.

Sin embargo, dotamos al problema de una nueva dimensión al razonar que existen infinitas definiciones posibles de una función en el espacio pseudo-Euclídeo que corresponden a la misma función de correlación en el espacio D . En efecto, el comportamiento de la función en la región de cuadrado negativo no tiene ninguna relevancia para un problema basado en el coste, ya que es imposible tener una configuración con distancias cuadráticas negativas.

Por tanto, la conclusión anterior tampoco invalida a la función exponencial como función de correlación para problemas basados en el coste, y nosotros la seguimos considerando una candidata prometedora. En este sentido existen infinitas funciones en el espacio pseudo-Euclídeo que se corresponden con la correlación exponencial en D . De hecho, probamos con algunas expresiones concretas e incluso conseguimos *corregir* la densidad espectral, aunque sólo en la región de cuadrado positivo.

Este capítulo abre innumerables líneas posibles de investigación futura. Algunas exigen profundizar en otras ramas de la matemática como el Análisis y la Teoría de la Medida, en el problema del cálculo del valor medio de una función sobre el hiperboloide d -dimensional. Otra línea abierta es la búsqueda de funciones definidas positivas a partir del cálculo de densidades espectrales

no negativas en el espacio pseudo-Euclídeo, utilizando las fórmulas integrales desarrolladas aquí.

También creemos necesario un cuestionamiento fundamental de esta línea de trabajo. ¿Se puede explotar de alguna manera el hecho de que las representaciones de problemas basados en el coste sólo adopten algunas de las posibles configuraciones en el espacio pseudo-Euclídeo? ¿Sería posible —y deseable— una definición alternativa de la transformada de Fourier definida en términos de la forma cuadrática pseudo-Euclídea, como propone [Strichartz \(1974\)](#)?

Por otra parte, en caso de que la representación pseudo-Euclídea demuestre tener alguna aplicación práctica, resultaría útil obtener fórmulas adicionales de representación, análogas a las de la Sección [6.2.2](#), para el caso en que \mathbf{X} sea degenerada y también para añadir más de un punto a la vez.

Capítulo 7

Reparametrización de la matriz de covarianzas

Hay una cantidad considerable de literatura dedicada a estimar la matriz de covarianzas Σ de un cierto vector aleatorio \mathbf{Y} , a partir de una muestra aleatoria $\mathbf{y}_1, \dots, \mathbf{y}_n$ de este vector. [Chiu et al. \(1996\)](#) modelan la matriz de covarianzas en términos de un conjunto de covariables, [Daniels y Pourahmadi \(2009\)](#) modelan las autocovarianzas de un vector de observaciones repetidas en el tiempo a partir de las autocorrelaciones parciales.

En todos los casos un problema importante consiste en garantizar que la matriz estimada sea definida positiva. Se han desarrollado diferentes estrategias, todas ellas basadas en alguna reparametrización o descomposición que permita estimar las componentes independientemente unas de otras, para luego reconstruir la matriz de covarianzas teniendo garantizada la definición positiva.

Nótese la similitud del problema con el que nos enfrentamos en el enfoque propuesto en esta Tesis. Aquí también se asume una estructura en la matriz de covarianzas, a saber, la dependencia sobre la distancia (basada en el coste). Y

el problema reside en garantizar que el modelo estimado garantice la definición positiva de la matriz.

Sin embargo hay también algunas diferencias importantes. En nuestro caso necesitamos no sólo garantizar que la matriz de covarianzas estimada para el vector de observaciones sea definida positiva, sino todas las posibles matrices generadas por la misma función de correlación, para todos los vectores posibles sobre todas las posibles localizaciones. Es decir, realmente buscamos un *modelo* definido positivo, que pueda haber generado una muestra como la observada.

Esto nos obliga a descartar algunos enfoques ingeniosos basados en la simulación, como el de [Barnard et al. \(2000\)](#), en el que se las arregla para obtener una muestra de la distribución posterior de la matriz de correlaciones asegurándose en cada paso que siempre es definida positiva. Aunque consigamos hacer inferencia sobre la estructura de correlaciones del vector observado, no tenemos ninguna información sobre el presunto modelo subyacente, ni garantía alguna de su existencia.

En el caso que nos ocupa, tenemos típicamente una única observación del vector \mathbf{Y} en lugar de una muestra de vectores. Pero a cambio, asumimos una dependencia funcional de las covarianzas sobre las distancias entre las localizaciones asociadas a las componentes de \mathbf{Y} . En cualquier caso, necesitamos garantizar que la estructura que estimamos a partir de los datos produzca una matriz definida positiva.

Parece natural revisar las diferentes parametrizaciones que se han utilizado en la literatura y comprobar si pueden aplicarse o adaptarse a nuestras circunstancias.

- [Barnard et al. \(2000\)](#) utilizan la descomposición en varianzas y correlaciones

$$\mathbf{\Sigma} = \mathbf{D}\mathbf{R}\mathbf{D},$$

donde la matriz diagonal \mathbf{D} contiene las desviaciones típicas y \mathbf{R} es la matriz de correlaciones. Desarrollan un muestreo de Gibbs en el cual simulan

una de las componentes de \mathbf{D} o \mathbf{R} por vez. La clave está en que es fácil determinar cuál es el rango de valores válido para una única correlación, suponiendo fijas las demás, para que la matriz \mathbf{R} sea siempre definida positiva. Este enfoque no nos sirve, porque no necesitamos únicamente estimar una matriz de correlaciones para un vector en concreto, sino un *modelo* válido para todas las posibles correlaciones entre cualquier conjunto de puntos.

- [Pourahmadi \(1999\)](#) utiliza una descomposición de Cholesky modificada sobre la matriz de precisiones

$$\boldsymbol{\Sigma}^{-1} = \mathbf{L}\mathbf{D}\mathbf{L}',$$

donde \mathbf{L} es una matriz triangular inferior con 1's en la diagonal y \mathbf{D} es diagonal con entradas positivas. Esta descomposición tiene el enorme atractivo de que elimina la restricción de definición positiva, en el sentido de que los elementos de \mathbf{L} y los logaritmos de los elementos de \mathbf{D} pueden tomar cualquier valor real independientemente unos de otros, y la composición seguirá siendo definida positiva. En efecto, para todo \mathbf{x}

$$\mathbf{x}'\boldsymbol{\Sigma}^{-1}\mathbf{x} = \mathbf{x}'\mathbf{L}\mathbf{D}\mathbf{L}'\mathbf{x} = (\mathbf{L}'\mathbf{x})'\mathbf{D}(\mathbf{L}'\mathbf{x}) \geq 0$$

pues \mathbf{D} es definida positiva. Además, los elementos de \mathbf{L} tienen una interpretación como los negativos de los coeficientes de regresión de cada elemento del vector aleatorio condicionado a los anteriores. Esto permite a los autores estimar las covarianzas a partir de unos diagramas análogos a los correlogramas de las series temporales, basados en una muestra de vectores aleatorios. Este enfoque tampoco nos sirve, porque se concentra en estimar la matriz de covarianzas de un vector aleatorio en concreto, del cual se dispone una muestra.

7.1 Descomposición de Cholesky

El enfoque tradicional en geoestadística consiste en modelar la correlación entre las localizaciones como una función de la distancia que las separa. Así se construye, a partir de las distancias entre observaciones, la matriz de covarianzas asociada a las mismas.

La matriz de covarianzas debe ser definida positiva y éste es el problema central de este trabajo. Pero toda matriz simétrica y definida positiva (SPD) \mathbf{A} admite una factorización en términos de una matriz triangular inferior \mathbf{L} : la *factorización de Cholesky*.

$$\mathbf{A} = \mathbf{L}\mathbf{L}',$$

donde \mathbf{L}' denota la matriz traspuesta de \mathbf{L} .

Si en lugar de expresar los elementos de la matriz de covarianzas como función de las distancias, modeláramos los elementos de la matriz triangular de Cholesky, podríamos construir la matriz de covarianzas correspondiente y sería automáticamente definida positiva.

Ahora bien, no podemos modelar estos elementos como función de las distancias directamente, porque entonces en general las covarianzas no serán necesariamente una función consistente de las distancias, como muestra el siguiente razonamiento.

Supongamos que modelamos los elementos \mathbf{L}_{ij} de la matriz de Cholesky como una función de los correspondientes elementos \mathbf{D}_{ij} de la matriz de distancias. La covarianza entre los puntos $i < j$ será entonces

$$\Sigma_{ij} = \sum_{k=1}^i \mathbf{L}_{ik} \mathbf{L}_{jk}.$$

Supongamos que ahora moviéramos un punto, por ejemplo el 1. Las distancias a los puntos i y j cambiarían, y cambiarían consecuentemente los elementos \mathbf{L}_{i1} y \mathbf{L}_{j1} y también su producto, salvo casos muy particulares. Por tanto cambiará

también la covarianza entre i y j , aunque la distancia entre ellos ha permanecido constante.

Por tanto, si queremos modelar las componentes de la matriz de Cholesky, debe realizarse de tal manera que las covarianzas sean una función de las distancias. Una posibilidad que hemos explorado es expresarlos como una función de una transformación de la matriz de distancias, análoga a la descomposición de Cholesky, como se muestra en el siguiente diagrama conmutativo.

$$\begin{array}{ccc} \mathbf{D} & \xrightarrow{f?} & \mathbf{\Sigma} \\ \downarrow T & & \downarrow \text{Cholesky} \\ T(\mathbf{D}) & \xrightarrow{g} & \mathbf{L}_{\Sigma} \end{array}$$

La idea consiste en pasar las distancias a un espacio *transformado*, donde se puedan modelar los elementos de \mathbf{L}_{Σ} como una función de estos elementos, para finalmente recuperar la matriz de covarianzas.

La transformación T no puede venir dada por la transformación de Cholesky porque \mathbf{D} no es en general SPD. Es natural pensar en la descomposición LU, que es una generalización de la descomposición de Cholesky aplicable a matrices de cualquier tipo.

Tendríamos que probar que esta estrategia de descomposición produce relaciones funcionales entre los elementos de las matrices transformadas.

Para ello, utilizamos el ejemplo de Curriero extendido con cuatro nuevos puntos (Fig 7.1).

Aplicando la función de correlación exponencial $\rho(h) = \exp(-h/\phi)$ a cada elemento de la matriz de distancias correspondiente, obtenemos una matriz de correlaciones definida positiva $\mathbf{\Sigma}$.

Esta matriz la descomponemos según Cholesky, obteniendo la matriz \mathbf{L}_{Σ} . Concretamente, para este ejemplo hemos utilizado la descomposición de Cholesky modificada $\mathbf{\Sigma} = \mathbf{L}_{\Sigma} \mathbf{d}_{\Sigma} \mathbf{L}_{\Sigma}'$ con el objetivo de estandarizar la matriz triangular poniendo los elementos de la diagonal en el vector \mathbf{d}_{Σ} .

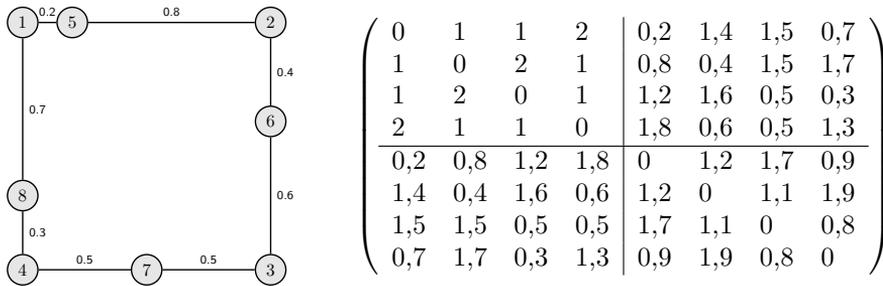


Figura 7.1. Extensión del ejemplo de Curriero y matriz de distancias \mathbf{D} .

Por otro lado, la matriz de distancias la descomponemos según la factorización PLU, ya que no tiene una descomposición LU pura. Aquí otra vez estandarizamos las diagonales de las matrices triangulares, escribiendo $\mathbf{D} = \mathbf{P}\mathbf{L}_D\mathbf{d}_D\mathbf{U}$.

Aquí nos encontramos con dos dificultades para establecer el vínculo elemento a elemento entre $T(\mathbf{D})$ y \mathbf{L}_Σ . En primer lugar, mientras que la matriz \mathbf{L}_Σ es única y caracteriza completamente la matriz de correlaciones, la descomposición PLU proporciona dos matrices triangulares distintas y no está claro cuál de ellas debería vincularse con \mathbf{L}_Σ . En segundo lugar, la presencia de la matriz \mathbf{P} implica un reordenamiento de las filas de \mathbf{L}_D (o, equivalentemente de \mathbf{D} , ya que $\mathbf{D} = \mathbf{P}\mathbf{L}_D\mathbf{U} \iff \mathbf{P}'\mathbf{D} = \mathbf{L}_D\mathbf{U}$), con lo cual se pierde la equivalencia entre los elementos correspondientes de \mathbf{L}_D y \mathbf{L}_Σ .

En la Figura 7.2 se representan los valores de los elementos (no nulos, fuera de la diagonal) de las matrices triangulares de la descomposición de la matriz \mathbf{D} contra los elementos correspondientes de la matriz triangular de Cholesky.

Se puede ver claramente no guardan una relación funcional, ni siquiera de forma aproximada. En esta representación hemos hecho caso omiso de la matriz \mathbf{P} , que ya habíamos anticipado que podía romper una eventual correspondencia. Pero si consideramos los elementos de la matriz $\mathbf{P}\mathbf{L}$, por ejemplo, ya no resulta tan evidente cuáles son los elementos correspondientes de la matriz \mathbf{L}_Σ .

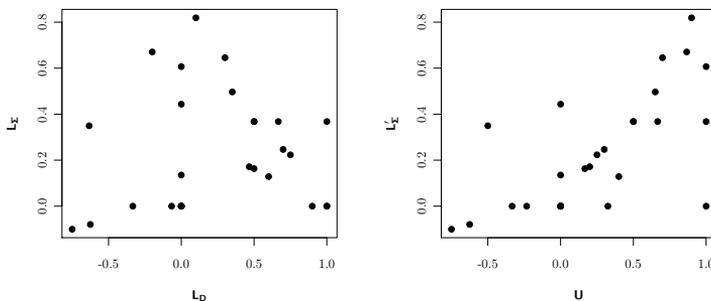


Figura 7.2. Relación empírica entre elementos correspondientes de las matrices triangulares.

7.2 Conclusiones

La Teoría Geoestadística se basa en asumir la existencia de una relación entre la distancia que separa dos localizaciones y la correlación entre los correspondientes valores de un campo aleatorio. Este vínculo se materializa en la función de correlación.

Como se ha visto en la introducción a esta Tesis, la utilización de la distancia Euclídea como medida de distancia entre localizaciones es inadecuada en regiones heterogéneas. Esto nos condujo a la propuesta de utilizar las distancias basadas en el coste introducidas en el Capítulo 3. Lo que a su vez, presenta problemas de orden teórico abordados en los Capítulos 5 y 6.

En este breve capítulo analizamos la posibilidad de un enfoque diferente del problema, sustituyendo la función de correlación por otra que vincule la distancia (eventualmente basada en el coste) con los elementos correspondientes de una reparametrización conveniente de la matriz de covarianzas. Si tal modelo tuviera una interpretación razonable, esta estrategia podría producir automáticamente matrices definidas positivas, garantizando así la validez del modelo.

Revisada la literatura relevante, encontramos que las estrategias de reparametrización generales no son adecuadas, en última instancia, porque no existe

un modelo subyacente que englobe todas las posibles matrices de covarianzas asociadas a cualquier conjunto de localizaciones seleccionado.

Con la estrategia de descomposición doble LU-Cholesky tampoco resulta claro entre qué elementos establecer una correspondencia funcional. El contraejemplo mostrado desmonta las posibilidades más naturales.

Damos así por zanjado un capítulo exploratorio en el que no vislumbramos otras líneas de trabajo interesantes.

Capítulo 8

Aproximaciones markovianas de campos Matérn

Durante la realización de esta Tesis se concretaron dos avances muy significativos en el campo de la inferencia Bayesiana en general y en la Geoestadística en particular. Aunque se venía gestando desde antes de 2005 (Rue y Held, 2005), en 2009 Rue et al. publican INLA (*integrated nested Laplace approximations*), una herramienta de software para la inferencia Bayesiana aproximada de modelos latentes Gaussianos. Al no utilizar métodos de simulación, INLA produce resultados muy rápidamente con un error de aproximación comparable al error de Monte Carlo. Desde entonces, el desarrollo de INLA no ha dejado de crecer y adquirir nuevas funcionalidades. En 2011, Lindgren et al. publican una serie de resultados que les permiten vincular ciertos campos Gaussianos de tipo Matérn con una aproximación markoviana discretizada (Gaussian Markov Random Fields, GMRFs). Esto a su vez les permite hacer inferencia de una manera mucho más eficiente (gracias a la propiedad de Markov), sin la necesidad de buscar una familia válida de funciones de covarianza, incluyendo problemas oscilatorios, no

estacionarios o sobre variedades diferenciales. Más aún, el hecho de trabajar con GMRFs hace posible la integración con INLA de manera natural.

Dos aspectos de este enfoque nos llamaron especialmente la atención y lo hacen relevante para nuestro problema. El primero es el hecho de que el *ajuste* del campo aleatorio no se hace en términos de una función de covarianza que tiene que ser válida *a priori* (nótese que éste es el meollo de la cuestión), sino que el campo se ajusta *localmente* (usando la propiedad de Markov) y gracias al vínculo con los campos continuos queda automáticamente probado que lo que se obtiene es una (buena) aproximación a un campo de tipo Matérn (válido). Esto parece ser cierto incluso cuando el campo está definido en una variedad diferencial más general, y éste es el segundo hecho de interés, ya que nuestras regiones con agujeros u otro tipo de irregularidades son efectivamente variedades diferenciales.

8.1 Vínculo entre GFs y GMRFs

Un campo Gaussiano (GF, por Gaussian Field) viene determinado por la función de covarianza (la función de media la asumimos nula, a estos efectos). Una función de covarianza que aparece naturalmente en diferentes campos científicos es la Matérn (Guttorp y Gneiting, 2006). Si denotamos por $\|\cdot\|$ a la distancia Euclídea en \mathbb{R}^d , la función de covarianza Matérn entre los localizaciones \mathbf{u} y $\mathbf{v} \in \mathbb{R}^d$ está definida como

$$k(\mathbf{u}, \mathbf{v}) = \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)} (\kappa\|\mathbf{v} - \mathbf{u}\|)^\nu K_\nu(\kappa\|\mathbf{v} - \mathbf{u}\|).$$

Aquí, K_ν es la función de Bessel modificada de segunda especie y orden $\nu > 0$, $\kappa > 0$ es un parámetro de escala (que determina el rango) y σ^2 es la varianza marginal.

Recientemente, las ecuaciones estocásticas en derivadas parciales (SPDEs, por *stochastic partial differential equations*) han demostrado ser de utilidad

para especificar de una manera alternativa ciertas clases de campos Gaussianos. Concretamente, un GF $x(\mathbf{u})$ con covarianza Matérn es una solución de la SPDE

$$(\kappa^2 - \Delta)^{\alpha/2} x(\mathbf{u}) = \mathcal{W}(\mathbf{u}), \quad \mathbf{u} \in \mathbb{R}^d, \alpha = \nu + d/2, \kappa > 0, \nu > 0,$$

donde $(\kappa^2 - \Delta)^{\alpha/2}$ es un pseudo-operador diferencial definido el [Lindgren et al. \(2011\)](#) a través de sus propiedades espectrales y \mathcal{W} es un proceso espacial de ruido blanco Gaussiano con varianza unitaria.

El uso de una SPDE permite la construcción de una aproximación GMRF a sus soluciones, lo que presenta importantes ventajas computacionales. Además, la introducción de parámetros aleatorios en los coeficientes de la SPDE nos permite considerarla como parte de un modelo jerárquico y ajustarlos en un esquema de inferencia Bayesiana.

Por su parte, un GMRF está definido por una matriz de precisión \mathbf{Q} , dispersa, que representa el comportamiento *condicional* del campo con respecto a una estructura de vecindades ([Rue y Held, 2005](#)). Nótese que \mathbf{Q} no es necesariamente la inversa de la matriz de covarianzas inducida sobre el grafo de vecindades por la función de covarianza ([Rue y Tjelmeland, 2002](#)).

El resultado principal en [Lindgren et al. \(2011\)](#) consiste básicamente en el cálculo de la matriz de precisión \mathbf{Q} , para ciertas funciones de covarianza en la familia Matérn, sobre cualquier triangulación de la región de interés, representando la estructura de vecindades asociada.

Una triangulación es una partición de la región en un conjunto de triángulos que se tocan a lo sumo en un lado completo o en un vértice, formando así un mosaico irregular. Tiene la ventaja de que se puede situar los vértices directamente sobre los puntos de observación, evitando así la aproximación habitual de la posición por el vértice más cercano en un mosaico regular. Además permite incrementar la resolución del mosaico sólo donde es necesario, ahorrando vértices en áreas con poca información.

Así por ejemplo, considerando una triangulación del cuadrado unidad (ver [Fig. 8.1](#), izquierda), podemos calcular la matriz de precisión \mathbf{Q} correspondiente

a una función de covarianza Matérn con parámetros prefijados. Invertiendo esta matriz, podemos comparar la covarianza correspondiente entre todos los vértices de la triangulación y un punto de referencia cualquiera, con la covarianza teórica que se ha pretendido aproximar.

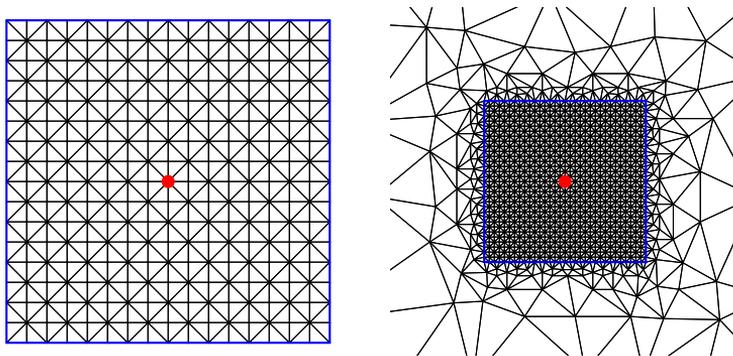


Figura 8.1. Triangulación regular del cuadrado unidad con un punto de referencia y una extensión menos densa para evitar el efecto de borde.

Sin embargo, la aproximación resulta ser muy mala, especialmente para rangos comparables al tamaño de la región (ver Fig. 8.2). Esto se debe a que las condiciones de borde utilizadas en la resolución de la SPDE introduce un *efecto de borde* que amplifica la varianza cerca de la frontera de la región. Empíricamente se ha visto que *cerca* significa aproximadamente una distancia equivalente a dos rangos del GF. Se está trabajando en la utilización de otras posibles condiciones de borde que no tendrían este efecto.

Por el momento, la solución práctica consiste en utilizar una triangulación *extendida* (ver Fig. 8.1, derecha), que lleve la frontera de la región suficientemente lejos del área de interés. En esta extensión nos podemos permitir utilizar una triangulación menos densa, para ahorrar tiempo de cálculo. Entonces sí observamos una correspondencia casi perfecta entre las covarianzas obtenidas por el método SPDE y las covarianzas teóricas (ver Fig. 8.3).

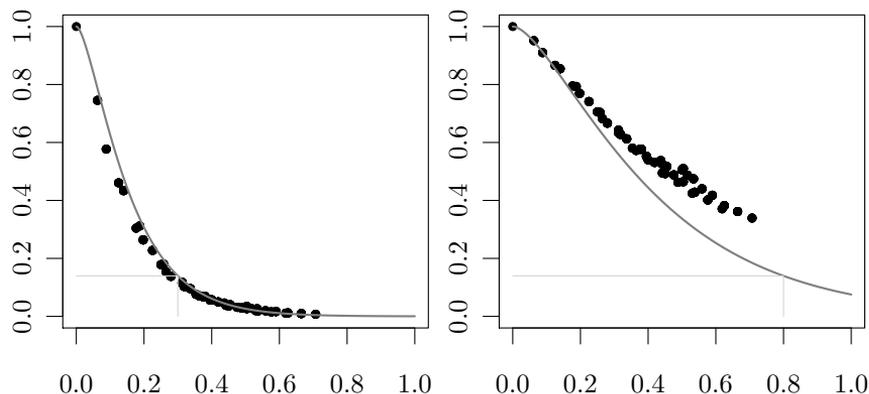


Figura 8.2. Correlaciones obtenidas mediante la aproximación SPDE comparadas con la función de correlación teórica. En este caso, una correlación tipo Matérn, con $\nu = 1$ y rangos 0,3 y 0,8 respectivamente. La triangulación utilizada se restringe al cuadrado, y el efecto de borde es el responsable de la discrepancia.

8.2 Estructura de correlaciones en la región de Curriero

Este vínculo entre GFs y GMRFs permite hacer la inferencia utilizando las propiedades locales de un GMRF, lo cual nos evita la necesidad de encontrar un modelo de covarianza global válido, lo que constituye el principal problema abordado en esta Tesis. Así pues, utilizando una triangulación adecuada de la región de trabajo (no necesariamente continua ni convexa), es posible ajustar un modelo con autocorrelación espacial que aproximará un campo Gaussiano continuo cuya validez estará garantizada por construcción.

En este tipo de regiones irregulares (por ejemplo, con *agujeros*) la correlación no puede ser una función de la distancia Euclídea. Nos ha interesado investigar cómo es la dependencia entre la correlación y la distancia cuando se utiliza este procedimiento en regiones irregulares. En particular, nos interesa saber si

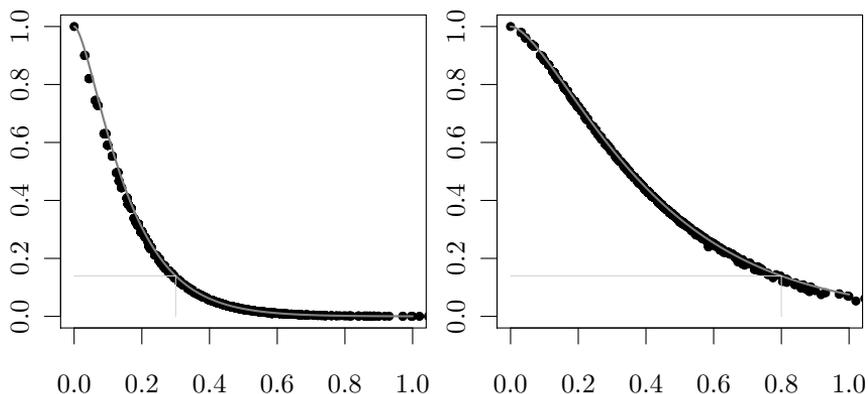


Figura 8.3. Correlaciones obtenidas mediante la aproximación SPDE comparadas con la función de correlación teórica. En este caso, una correlación tipo Matérn, con $\nu = 1$ y rangos 0,3 y 0,8 respectivamente. La utilización de una triangulación extendida corrige el efecto de borde.

la estructura de correlaciones obtenidas por el método SPDE puede tener una relación (aproximadamente) funcional con la distancia basada en el coste.

Para comprobar las propiedades de este procedimiento, nos planteamos la región de Curriero como caso de pruebas. En esta región, definida como \mathbb{R}^2 menos el cuadrado unidad, establecemos una triangulación y calculamos la correspondiente matriz de precisión \mathbf{Q} asociada a un modelo Matérn con parámetros prefijados. También definiremos un punto de referencia respecto del cual analizaremos las correlaciones con todos los demás vértices de la triangulación (ver Fig. 8.4).

En la Figura 8.5 se comprueba que, en efecto, no existe una relación funcional entre la correlación y la distancia Euclídea en una región con estas características.

Las distancias basadas en el coste son las longitudes de las curvas más cortas que conectan dos puntos en la región. En la Figura 8.6 se puede ver, en trazo negro, la curva de mínimo coste que conecta un punto de la región con el punto de referencia. En este caso, la distancia basada en el coste es de 0,67 unidades, mientras que la distancia Euclídea es de 0,64. El mapa de colores

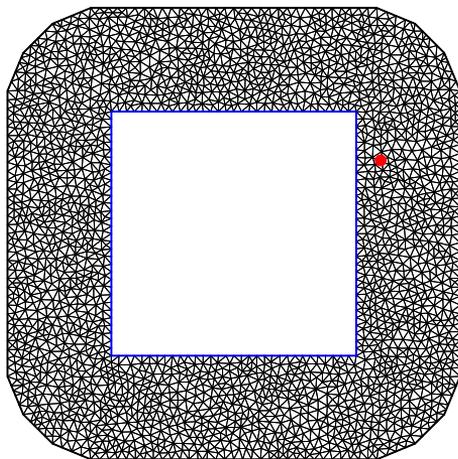


Figura 8.4. Triangulación de la región de Curriero y punto de referencia.

indica la distancia basada en el coste entre cualquier punto de la región y el punto de referencia. De esta manera, podemos representar las correlaciones entre todos los vértices de la triangulación y el punto de referencia con respecto a las correspondientes distancias basadas en el coste (ver Fig. 8.7)

Se puede apreciar una dispersión algo menor en la nube de puntos alrededor de la curva teórica, lo que indica que la distancia basada en el coste refleja algo mejor que la Euclídea la estructura de correlaciones generada por el método SPDE.

Sin embargo la relación tampoco es funcional, claramente. Podríamos pensar que el efecto de borde es el responsable de la discrepancia, pero no lo es. Aunque este punto no es inmediato de resolver.

En efecto, en este tipo de regiones, utilizar un triangulación extendida como en la Sección 8.1 implica *conectar* puntos a través del pretendido agujero (ver Fig. 8.8), eliminando las discontinuidades de la región. El efecto sobre la estructura de correlaciones del método SPDE es dotarlas de una estructura Euclídea,

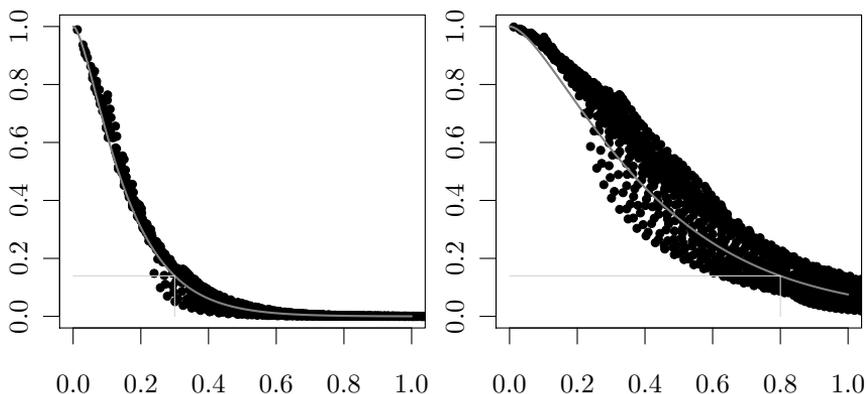


Figura 8.5. Correlaciones entre los vértices de la triangulación y el punto de referencia en función de la **distancia Euclídea** para modelos Matérn con rangos 0,3 y 0,8 respectivamente. La línea continua representa la función de correlación teórica correspondiente.

obteniendo ajustes casi perfectos del modelo teórico Euclídeo, como los de la Figura 8.3.

La manera práctica que encontramos de eliminar o mitigar el efecto de borde consiste en escapar del plano, definiendo la triangulación sobre una superficie tal que los puntos interiores del cuadrado unidad tienen una cierta elevación y forman una especie de *pirámide*. El modelo SPDE se puede definir sobre una superficie de este tipo y el efecto que esperamos obtener es similar al obtenido con la región discontinua. A saber, aumentar la distancia efectiva entre dos puntos separados por el *hueco* interior.

La estructura de correlaciones obtenida, en función de la distancia basada en el coste, se muestra en la Figura 8.9. Aunque ha mejorado algo el ajuste, no queda lugar a dudas que la distancia basada en el coste no reproduce la estructura de correlaciones obtenidas por el método SPDE, sino de una manera muy groseramente aproximada.

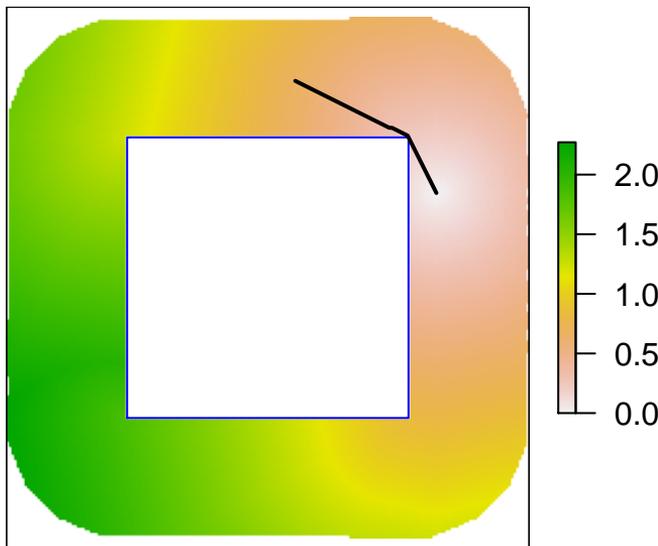


Figura 8.6. Mapa de distancias acumuladas al punto de referencia. En trazo negro se muestra un camino de mínimo coste.

8.3 Conclusiones

En este capítulo hemos podido comprobar, con un procedimiento alternativo, cómo la clásica estructura de correlaciones basada en distancias Euclídeas no resulta adecuada en regiones no homogéneas y, muy en particular, en regiones discontinuas.

Sin embargo, también hemos comprobado que la estructura de correlaciones inducida por el método basado en SPDE tampoco responde a una función de nuestra definición de distancia basada en el coste. Resultaría interesante estudiar qué tipo de estructura de distancias es inducido por este método, asumiendo una relación funcional entre correlaciones y distancias.

Dado el contraejemplo, no parece posible establecer ningún vínculo teórico entre el modelo SPDE y el modelo definido a través de la distancia basada en

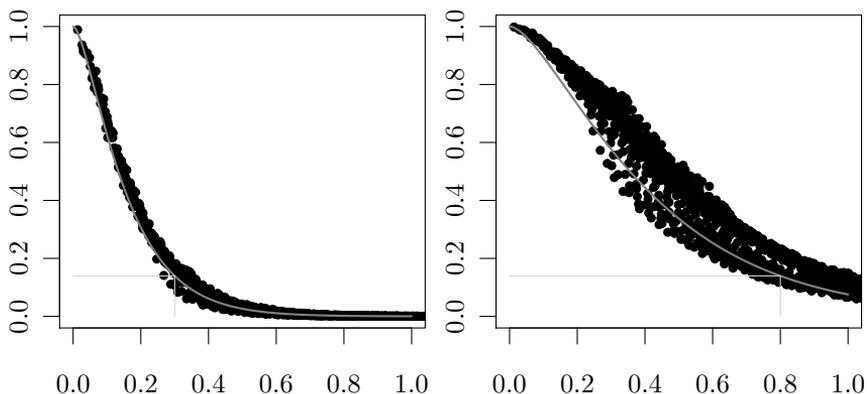


Figura 8.7. Correlaciones entre los vértices de la triangulación y el punto de referencia en función de la **distancia basada en el coste** para modelos Matérn con rangos 0,3 y 0,8 respectivamente. La línea continua representa la función de correlación teórica correspondiente.

el coste. Este hubiera sido nuestro objetivo, utilizando esta herramienta como *punteo* teórico que nos permitiera validar la utilización de las distancias basadas en el coste.

Es importante recalcar que este procedimiento efectivamente resuelve el problema de trabajar en regiones discontinuas, aunque el efecto del borde es un aspecto a pulir y en el que nos consta que se está trabajando. Además, tiene la virtud de establecer una plataforma estándar para la especificación de ciertos tipos de modelos jerárquicos Bayesianos con variable respuesta generalizada (en la familia exponencial) y en los que se pueden incorporar covariables con facilidad.

Por otro lado, el método SPDE no resulta tan general como el basado en el coste, ya que las distancias basadas en el coste permiten modelar todo tipo de irregularidades mediante la superficie de coste, y no solamente las discontinuidades.

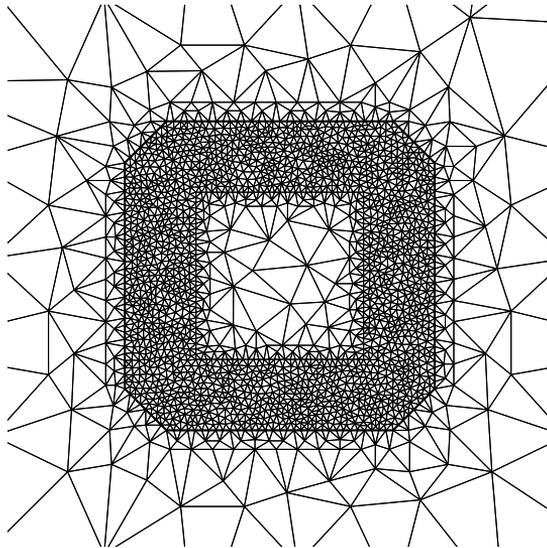


Figura 8.8. Una triangulación extendida de la región de Curriero tiene el efecto indeseado de eliminar las discontinuidades de la región.

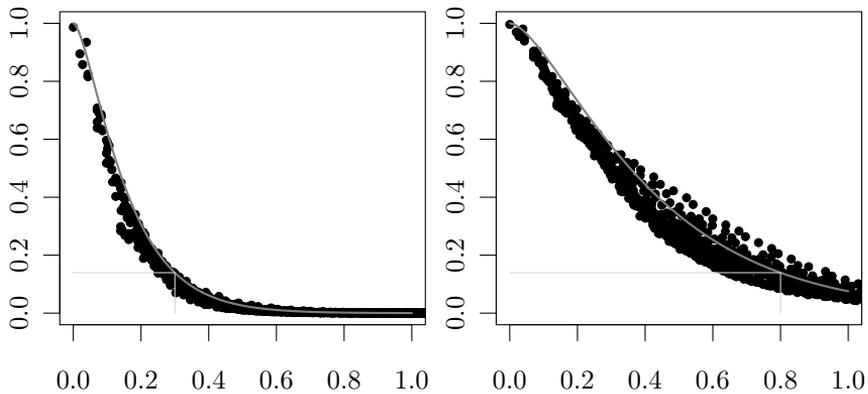


Figura 8.9. Correlaciones entre los vértices de la triangulación y el punto de referencia en función de la **distancia basada en el coste** para modelos Matérn con rangos 0,3 y 0,8 respectivamente, **sin efecto de borde**. La línea continua representa la función de correlación teórica correspondiente.

Capítulo 9

Conclusiones y líneas futuras de trabajo

9.1 Conclusiones generales

En esta Tesis se aborda el problema de la predicción geoestadística en regiones heterogéneas, de manera tal que se pueden distinguir dos contribuciones principales.

Por una parte (Caps. 3 y 4), hemos dado una respuesta práctica y aplicada al problema desarrollando, implementando y aplicando una metodología general para este tipo de escenarios, que denominamos *basada en el coste*.

Por otra (Cap. 5), y motivados por lo anterior, analizamos rigurosamente el problema matemático de la búsqueda de funciones definidas positivas en problemas basados en el coste. En particular (Cap. 6), demostramos que este tipo de problemas siempre pueden representarse en un espacio pseudo-Euclídeo, donde buscamos caracterizaciones de las funciones definidas positivas *isotrópicas* (i.e., que sólo dependen de la distancia).

Además (Caps. 7 y 8), exploramos algunos enfoques alternativos al problema, que no utilizan la metodología basada en el coste.

9.2 Conclusiones específicas

Más concretamente, identificamos en primer lugar (Cap. 3) el problema en el caso concreto de la elaboración de mapas acústicos en entornos urbanos, que utilizamos como piloto de pruebas, pero que inmediatamente generalizamos a toda suerte de situaciones de predicción espacial en la cual la región de trabajo presente irregularidades importantes.

En el mismo capítulo proponemos la metodología *basada en el coste*, para tratar este tipo tan general de problemas. La definimos rigurosamente, analizamos sus propiedades y observamos que generaliza varias técnicas utilizadas en situaciones más específicas. También la implementamos en una plataforma informática adecuada (Cap. 4); a saber, el Sistema de Información Geográfica GRASS y el paquete geoestadístico *geoR*. Esto involucra fundamentalmente dos operaciones no triviales: el cálculo de las distancias basadas en el coste y la adaptación de los algoritmos geoestadísticos para utilizar distancias basadas en el coste en lugar de distancias Euclídeas. Finalmente, utilizando las herramientas programadas, aplicamos la metodología al proyecto piloto presentando los resultados y comparándolos con los obtenidos con la metodología clásica.

La metodología basada en el coste produce predicciones válidas, siempre que se compruebe la definición positiva de la matriz de covarianzas de las observaciones. Pero no hay garantías de que se pueda utilizar en cualquier situación, ni para cualquier banco de datos. Tampoco se puede garantizar la validez del campo aleatorio subyacente en el modelo estadístico. En suma, la metodología basada en el coste proporciona un interpolador lineal (de hecho, el *mejor predictor lineal insesgado*, o BLUP por sus siglas en inglés) con valoración del error de predicción, pero no necesariamente de un modelo estadístico válido.

La validez del modelo estadístico se reduce a la condición de definición positiva de la función de correlación del modelo, cuando las distancias utilizadas son

basadas en el coste. Esta condición exige que para cualquier número n , conjunto de localizaciones $\mathbf{s}_1, \dots, \mathbf{s}_n$ y conjunto de coeficientes complejos $\alpha_1, \dots, \alpha_n$, la función R verifique la relación

$$\sum_{i,1=1}^n \bar{\alpha}_i \alpha_j R(\mathfrak{d}(s_i, s_j)),$$

donde \mathfrak{d} representa la distancia basada en el coste entre sus argumentos.

Mientras que en el espacio Euclídeo existe una caracterización de (todas) las funciones definidas positivas, en el *espacio basado en el coste* es un problema que hemos analizado desde diferentes puntos de vista (Cap. 5). El espacio basado en el coste se puede describir adecuadamente como una variedad Riemanniana cuya métrica depende de la superficie de coste asociada al problema. Además, la métrica Riemanniana induce naturalmente una función de distancia que coincide precisamente con la distancia basada en el coste. Por lo tanto, se trata de investigar la condición de definición positiva de las funciones es esta variedad Riemanniana.

Una estrategia para abordar este problema es la inmersión del espacio basado en el coste en otro espacio con estructuras matemáticas adicionales. Aunque hemos demostrado (Prop. 6.4) que la inmersión en un espacio Euclídeo o Hilbert es inviable en general, es posible realizar otras inmersiones como la de Kuratowski, que sumerge el espacio basado en el coste en un subconjunto de un Álgebra de Banach.

De las diversas posibilidades, la que juzgamos más interesante, y en la que invertimos buena parte de nuestro trabajo, fue la inmersión en un espacio pseudo-Euclídeo $E_{(k,d-k)}$ de dimensión d e índice k (Cap. 6). Éste, es una generalización del espacio Euclídeo en la cual el producto interno se sustituye por una forma bilineal simétrica en la que la condición de definición positiva es sustituida por la de no-degeneración (Def. 6.5). Esto produce un espacio con k dimensiones *normales*, con contribuciones positivas a las distancias y $d - k$ dimensiones que contribuyen negativamente; lo que divide el espacio en dos regiones, separadas

por un (hiper)cono. Una región contiene puntos con distancia cuadrática positiva respecto del origen y la otra está formada por puntos con distancia cuadrática negativa.

Inspirados en la técnica conocida como Multidimensional Scaling (MDS), que proporciona representaciones aproximadas en espacios Euclídeos, demostramos (Teo. 6.6) que siempre es posible encontrar una representación pseudo-Euclídea de cualquier problema basado en el coste. Lo recíproco, en cambio, no es cierto. Es decir, no todo conjunto de puntos en el espacio pseudo-Euclídeo responde a una representación de alguna configuración basada en el coste. Esto se manifiesta claramente en la desigualdad triangular, propiedad satisfecha por la métrica basada en el coste, pero no por la premétrica simétrica inducida en el espacio pseudo-Euclídeo.

Esto parecería indicar que la condición de definición positiva es más exigente en el espacio pseudo-Euclídeo, que en el métrico D . Sin embargo, hemos argumentado que el conjunto $\mathcal{P}(D)$ de funciones multivariantes definidas positivas en el espacio D debe ser el mismo que en el espacio pseudo-Euclídeo (y que en el Euclídeo). La complejidad sobreviene cuando nos ocupamos de la subfamilia de funciones isotrópicas, que es la que verdaderamente nos interesa desde el punto de vista de las aplicaciones. La obtención de la representación espectral de una función isotrópica exige calcular el valor medio de una función exponencial compleja sobre una superficie hiperboloide, de extensión infinita. Para esto fue necesario desarrollar un sistema de coordenadas (Ap. F) adaptado al producto pseudo-Euclídeo.

El cálculo del valor medio de una función sólo puede definirse con seguridad en un dominio de medida finita. En caso contrario, el resultado puede ser diferente dependiendo de cómo se lleve a cabo el proceso de límite. Es habitual utilizar el *Valor Principal de Cauchy* como criterio para asignar valores a integrales impropias de este tipo. Aún así, sólo hemos sido capaces de obtener resultados parciales en el cálculo de la caracterización de funciones definidas positivas isotrópicas en general. También proporcionamos resultados parciales en algunos casos particulares, como la densidad espectral de la función constante, las

fórmulas integrales para la densidad espectral de una función isotrópica general en el espacio $E_{(2,1)}$ y la densidad espectral de la función exponencial en el mismo espacio.

De esta manera, llegamos a la inesperada conclusión de que la función exponencial *no es* definida positiva en $E_{(2,1)}$ y, por tanto, no está legitimada, en principio, como función de correlación para un problema basado en el coste. Sin embargo, dotamos al problema de una nueva dimensión al razonar que existen infinitas definiciones posibles de una función en el espacio pseudo-Euclídeo que corresponden a la misma función de correlación en el espacio D . En efecto, el comportamiento de la función en la región de cuadrado negativo no tiene ninguna relevancia para un problema basado en el coste, ya que es imposible tener una configuración con distancias cuadráticas negativas. Por tanto, la conclusión anterior no invalida a la función exponencial como función de correlación para problemas basados en el coste y nosotros la seguimos considerando una candidata prometedora. En este sentido, existen infinitas funciones en el espacio pseudo-Euclídeo que se corresponden con la correlación exponencial en D . De hecho, probamos con algunas expresiones concretas consiguiendo *corregir* la densidad espectral, aunque sólo en la región de cuadrado positivo.

Un enfoque diferente que hemos explorado, ha sido la sustitución de la función de correlación por otra que vincule la distancia (eventualmente basada en el coste) con los elementos correspondientes de una reparametrización conveniente de la matriz de covarianzas (Cap. 7). Si tal modelo tuviera una interpretación razonable, esta estrategia podría producir automáticamente matrices definidas positivas, garantizando así la validez del modelo.

Sin embargo, las estrategias de reparametrización encontradas en la literatura se enfocan todas a obtener matrices de covarianza definidas positivas, pero carecen de un modelo subyacente que garantice la definición positiva para cualquier conjunto potencial de observaciones.

También exploramos una estrategia de descomposición doble LU-Cholesky con la cual tampoco resulta claro entre qué elementos establecer una correspondencia funcional. Con un contraejemplo se descartan las posibilidades más naturales.

Por último, consideramos imprescindible analizar una nueva metodología desarrollada durante la realización de esta Tesis y estudiar su vinculación con la metodología basada en el coste. Se trata de la aproximación mediante campos markovianos (Gaussian Markov Random Fields, GMRFs) de campos aleatorios Gaussianos con covarianza de tipo Matérn (Lindgren et al., 2011), a través de la solución numérica de una Ecuación Diferencial en Derivadas Parciales (Stochastic Partial Differential Equation, SPDE). Esto permite realizar inferencia de una manera mucho más eficiente (gracias a la propiedad de Markov), sin la necesidad de buscar una familia válida de funciones de covarianza. En cambio, el campo se ajusta *localmente* y gracias al vínculo con los campos continuos, queda automáticamente probado que lo que se obtiene es una (buena) aproximación a un campo aleatorio de tipo Matérn (válido). Esto es cierto incluso cuando el campo está definido en una variedad diferencial más general, como regiones con discontinuidades.

Utilizando esta metodología hemos vuelto a confirmar que la clásica estructura de correlaciones basada en distancias Euclídeas no resulta adecuada en regiones no homogéneas y, muy en particular, en regiones discontinuas.

Por otra parte, también hemos comprobado que la estructura de correlaciones inducida por el método basado en SPDE no es equivalente a la inducida por la metodología basada en el coste. Dado el contraejemplo, no parece posible establecer ningún vínculo teórico entre el modelo SPDE y el modelo definido a través de la distancia basada en el coste. De otra manera, habríamos utilizando esta metodología como puente teórico para validar la utilización de las distancias basadas en el coste.

Es importante recalcar que este procedimiento efectivamente resuelve el problema de trabajar en regiones discontinuas. Además, tiene la virtud de establecer una plataforma estándar para la especificación de ciertos tipos de modelos jerárquicos Bayesianos con variable respuesta generalizada (en la familia exponencial) y en los que se pueden incorporar covariables con facilidad.

Sin embargo, el método SPDE no resulta tan general como el basado en el coste, ya que las distancias basadas en el coste permiten modelar todo tipo de

irregularidades mediante la superficie de coste y no solamente las discontinuidades.

9.3 Líneas abiertas de investigación

En lo que se refiere al problema particular de los mapas acústicos en entornos urbanos, creemos que sería muy ventajoso utilizar los resultados de los métodos de cálculo como predictores de la media del proceso. La aportación Bayesiana al campo de la *Calibración de Modelos Complejos y Data Assimilation* ha sido muy fructífera en los últimos años. Ver [Bhat et al. \(2010\)](#) como ejemplo de implementación de este tipo de metodologías a modelos con respuesta espacial, como es el caso de los modelos de difusión de ruido. Esto, junto con la herramienta de geoestadística basada en el coste, completaría un sistema integrado de predicción que combinaría de forma armoniosa las distintas herramientas de análisis y predicción disponibles, aprovechando lo mejor de cada una y propiciando una sinergia sin lugar a dudas beneficiosa desde el punto de vista de la predicción. Lamentablemente, la falta de disponibilidad de herramientas de simulación libres (esto es, con una licencia que permita la utilización y modificación del código fuente) entorpece la investigación en este sentido.

En cuanto a la caracterización de las funciones definidas positivas para problemas basados en el coste, la línea de trabajo que pasa por la inmersión de D en un espacio de Banach cuenta con la interesante generalización del Teorema de Bochner sobre álgebras de Banach ([Rudin, 1991](#), Teo. 11.32). Tenemos además la caracterización de [Berg et al. \(1984, Teo. 5§2.1\)](#) sobre funciones definidas positivas isotrópicas en espacios de Banach. Aunque también sabemos ([Berg et al., 1984, Teo. 5§2.3](#)) que éstas son triviales en el caso del Banach $\mathcal{C}[0, 1]$, precisamente uno de los espacios donde sabemos representar D .

Éste es un camino aún por recorrer, del cual los teoremas y resultados comentados son hitos que es imprescindible tener en cuenta y aprovechar. En su caso, es necesario un estudio teórico más riguroso de la Teoría Espectral en álgebras de Banach que el que hemos esbozado aquí.

Por otra parte, y evitando las inmersiones, sería interesante considerar el grupo de movimientos rígidos del espacio vectorial \mathbb{R}^d actuando sobre la variedad D . Una *traslación* en este contexto se efectuaría recorriendo la geodésica que pasa por el punto de origen, a lo largo de una distancia determinada por la longitud del vector traslación. Esta estructura adicional permitiría introducir el concepto de estacionariedad en el espacio D , habilitando la utilización de los resultados de [Bochner \(1941\)](#) sobre la representación espectral de funciones definidas positivas estacionarias. La caracterización de la isotropía, sin embargo, quedaría por resolver. Una posibilidad natural podría ser la medida de Hausdorff ([Federer, 1978](#)), que generaliza la de Lebesgue en \mathbb{R}^d , permitiendo así considerar una transformación de Fourier generalizada ([Lindgren et al., 2011](#)). En este particular, es interesante el trabajo de [Strichartz \(1974\)](#) sobre las transformadas de Fourier de funciones invariantes ante la acción de un grupo (de rotaciones). Aún así, consideramos que ésta es la línea de trabajo más prístina, directa y promisoría.

Por vía de la representación pseudo-Euclídea también se abren innumerables líneas posibles de investigación futura. Algunas, como el problema del cálculo del valor medio de una función sobre el hiperboloide d -dimensional, exigen profundizar en otras ramas de la matemática como el Análisis y la Teoría de la Medida.

Otra línea abierta es la búsqueda directa de funciones definidas positivas a partir del cálculo de densidades espectrales no negativas en el espacio pseudo-Euclídeo, utilizando las Fórmulas integrales (6.12).

También creemos necesario un cuestionamiento fundamental de esta línea de trabajo. Nos preguntamos, por ejemplo, si es posible explotar de alguna manera el hecho de que las representaciones de problemas basados en el coste sólo adopten algunas de las posibles configuraciones en el espacio pseudo-Euclídeo. O, alternativamente, si es posible *filtrar* las frecuencias de cuadrado negativo en la construcción de la función de correlación. Si así fuera, tendríamos una expresión definida positiva de la función exponencial.

Por otra parte, en caso de que la representación pseudo-Euclídea demuestre tener alguna aplicación práctica, resultaría útil obtener fórmulas adicionales de representación, análogas a las de la Sección 6.2.2, para el caso en que la matriz de coordenadas fuera degenerada y también para añadir más de un punto a la vez.

Hay otros enfoques posibles del problema que por cuestiones de tiempo se han postergado. Uno de ellos consiste en la comprobación numérica al estilo *fuerza bruta* de la definición positiva de una función candidata. El método consiste en simular conjuntos de distancias basadas en el coste, comprobando numéricamente la definición positiva de la matriz de correlaciones resultante de aplicar la función en cuestión. Para tener una mínima garantía de validez, sería necesario asegurar que para un número n de puntos simulados, se generan suficientes conjuntos de $n(n-1)/2$ distancias como para cubrir razonablemente el espacio de posibilidades. Si se piensa cada conjunto de distancias como un punto en el espacio $\mathbb{R}^{n(n-1)/2}$, la región de puntos compatible con una métrica es un cono en el cual habría que tomar una muestra suficientemente densa, hasta cierto límite razonable. Aunque se concentra principalmente en el cono de las distancias Euclídeas, incluido en el que nos interesa, Dattorro (2010) proporciona un material de mucha utilidad en este sentido. Krivoruchko y Gribov (2002, p.8) afirman haber demostrado por este método que la función exponencial no es definida positiva en general, pero proporcionan pocos detalles acerca de qué tipo de distancias utilizaron y el código fuente del experimento no está disponible como para ser reproducido por otros investigadores.

Otro enfoque, de similar empirismo, consiste en estudiar, para una superficie de coste determinada, cuál es el conjunto de funciones definidas positivas, dentro de alguna familia suficientemente flexible. Para esto sería necesario construir alguna familia a partir de alguna base de funciones, con sendos parámetros, como las que paradójicamente se denominan *no paramétricas* . Además, simular múltiples configuraciones de puntos en dicha región y determinar las funciones que producen matrices de correlación definidas positivas en todas ellas. Este procedimiento, que restringe sus resultados a un problema específico, podría

ser utilizado para ganar *intuición* acerca de las posibles candidatas a funciones válidas.

La predicción geoestadística en regiones heterogéneas es un problema complejo, de altísimo interés aplicado y objeto de un intenso trabajo de investigación en todo el mundo. Está lejos de ser zanjado, aunque los avances en algunas situaciones específicas son notables: problemas en redes de ríos ([Ver-Hoef et al., 2006](#)), o en regiones discontinuas ([Lindgren et al., 2011](#)), problemas con anisotropía local conocida ([Boisvert et al., 2009](#)) o que responden a leyes físicas conocidas ([Christakos y Serre, 2000](#)), etc. Creemos y esperamos que esta Tesis sea una modesta contribución al camino que hacemos al andar todos los que hemos procurado el avance científico en esta materia.

Apéndice

Apéndice A

Deducción de las ecuaciones del Kriging

Los datos $\mathbf{Z} = \{Z(\mathbf{s}_1), \dots, Z(\mathbf{s}_n)\}$ se suponen una realización parcial de un proceso estocástico

$$\{Z(\mathbf{s}) : \mathbf{s} \in D \subseteq \mathbb{R}^n\},$$

que se descompone en una estructura determinista $\mu(\cdot) \equiv \mathbb{E}[Z(\cdot)]$ y un proceso de error espacialmente correlacionado $\delta(\cdot)$ con media nula. Sobre estas componentes se hacen hipótesis adicionales según el caso.

El predictor Kriging es el predictor lineal insesgado óptimo en el sentido que minimiza la varianza del error de predicción (Kriging variance, prediction variance), o lo que es lo mismo, el error cuadrático medio.

Sea \mathbf{s}_0 la localización donde se desea predecir, y $Z(\mathbf{s}_0)$ el verdadero valor, no observado. Entonces,

$$\hat{Z}(\mathbf{s}_0) = \sum_{i=1}^n \lambda_i Z(\mathbf{s}_i),$$

tal que se verifican:

$$i) \text{ (insesgadez) } \mathbb{E} \left[\hat{Z}(\mathbf{s}_0) \right] = \mathbb{E} [Z(\mathbf{s}_0)],$$

$$ii) \text{ (optimalidad) } \min \mathbb{V} \left[\hat{Z}(\mathbf{s}_0) - Z(\mathbf{s}_0) \right].$$

En primer lugar, notar que si el proceso tiene media constante μ (y simplificando la notación)

$$\mathbb{E} \left[\hat{Z} \right] = \mathbb{E} [Z_0] \iff \mathbb{E} \left[\sum_{i=1}^n \lambda_i Z_i - Z_0 \right] = 0 \iff \left(\sum_{i=1}^n \lambda_i - 1 \right) \mu = 0.$$

De modo que para que el predictor sea insesgado, o bien la media del proceso es $\mu = 0$, o bien los coeficientes suman 1. El primer caso da lugar al *Kriging Simple*. Notar que si la media del proceso es una constante conocida μ entonces $Z - \mu$ es un proceso de media conocida cero. Así, el Kriging Simple se puede aplicar a procesos con media constante conocida.

Más habitual es utilizar el *Kriging Ordinario*, en el cual se asume que el proceso tiene una media constante μ , pero desconocida, y por tanto es necesario trabajar con la restricción $\sum_{i=1}^n \lambda_i = 1$.

En segundo lugar, notar que minimizar la varianza del error es lo mismo que minimizar el error cuadrático medio, bajo la restricción de insesgadez. Esta cantidad se denomina *varianza de predicción* o *Kriging variance*:

$$\mathbb{V} \left[\hat{Z} - Z_0 \right] = \mathbb{E} \left[(\hat{Z} - Z_0)^2 \right] - \underbrace{\left(\mathbb{E} \left[\hat{Z} - Z_0 \right] \right)^2}_{=0} \equiv \sigma_k^2.$$

Si el proceso tiene estacionariedad de segundo orden, entonces en particular es homocedástico ($\mathbb{V} [Z(\mathbf{s})] = C(\mathbf{0}) = \sigma^2$ (cte.) $\forall \mathbf{s} \in D$, donde $C(\cdot)$ es la función de covarianza), y además se puede trabajar con la estructura de covarianzas.

Entonces,

$$\begin{aligned}\sigma_k^2 &= \mathbb{V} [\hat{Z} - Z_0] = \mathbb{V} \left[\sum_{i=1}^n \lambda_i Z_i \right] - 2\mathbb{C} \left[\sum_{i=1}^n \lambda_i Z_i \right] Z_0 + \mathbb{V} [Z_0] \\ &= \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \mathbb{C} [Z_i] Z_j - 2 \sum_{i=1}^n \lambda_i \mathbb{C} [Z_i] Z_0 + \sigma^2.\end{aligned}$$

Matricialmente se expresa más sencillamente:

$$\boxed{\sigma_k^2 = \boldsymbol{\lambda}' \boldsymbol{\Sigma} \boldsymbol{\lambda} - 2\boldsymbol{\lambda}' \mathbf{c} + \sigma^2.} \quad (\text{A.1})$$

A.1 Kriging Simple

Se asumen las siguientes hipótesis:

- $\mu(\mathbf{s}) = \mu$ (cte. conocida) (sin pérdida de generalidad, $\mu = 0$)
- estacionariedad de segundo orden

Se minimiza la varianza de predicción por derivación respecto del vector de incógnitas $\boldsymbol{\lambda}$.

$$\nabla(\sigma_k^2) = 0 \iff (\boldsymbol{\Sigma} + \boldsymbol{\Sigma}') \boldsymbol{\lambda} - 2\mathbf{c} = 0,$$

donde se han utilizado las propiedades de derivación vectorial $\nabla(\mathbf{a}\mathbf{x}) = \mathbf{a}$,¹ y $\nabla(\mathbf{x}'\mathbf{A}\mathbf{x}) = (\mathbf{A} + \mathbf{A}')\mathbf{x}$.² Y como la matriz $\boldsymbol{\Sigma}$ es simétrica,

$$\boldsymbol{\lambda} = \boldsymbol{\Sigma}^{-1} \mathbf{c}.$$

Por tanto, el predictor Kriging resulta ($\boldsymbol{\Sigma}^{-1}$ simétrica),

$$\hat{Z} = \boldsymbol{\lambda}' \mathbf{Z} = \mathbf{c}' \boldsymbol{\Sigma}^{-1} \mathbf{Z},$$

¹ $\nabla(\mathbf{a}\mathbf{x}) = \left(\frac{\partial \mathbf{a}\mathbf{x}}{\partial x_1}, \dots, \frac{\partial \mathbf{a}\mathbf{x}}{\partial x_n} \right) = (a_1, \dots, a_n) = \mathbf{a}$.

² $\nabla(\mathbf{x}'\mathbf{A}\mathbf{x}) = \nabla(\sum_i \sum_j x_i x_j a_{ij}) = (\sum_i \lambda_i a_{i1} + \sum_j \lambda_j a_{1j}, \dots, \sum_i \lambda_i a_{in} + \sum_j \lambda_j a_{nj}) = (\sum_i \lambda_i a_{i1}, \dots, \sum_i \lambda_i a_{in}) + (\sum_j \lambda_j a_{1j}, \dots, \sum_j \lambda_j a_{nj}) = (\boldsymbol{\lambda}'\mathbf{A})' + \mathbf{A}\boldsymbol{\lambda} = (\mathbf{A} + \mathbf{A}') \cdot \mathbf{x}$.

con varianza de predicción

$$\begin{aligned}\sigma_k^2 &= \boldsymbol{\lambda}' \boldsymbol{\Sigma} (\boldsymbol{\Sigma}^{-1} \mathbf{c}) - 2\boldsymbol{\lambda}' \mathbf{c} + \sigma^2 = \sigma^2 - \boldsymbol{\lambda}' \mathbf{c} \\ \sigma_k^2 &= \sigma^2 - \mathbf{c}' \boldsymbol{\Sigma}^{-1} \mathbf{c}.\end{aligned}$$

Como $\boldsymbol{\Sigma}$ y su inversa son definidas positivas, se desprende que esta varianza será siempre menor o igual que la varianza del proceso.

A.2 Kriging Ordinario

Se asumen las siguientes hipótesis:

- $\mu(\mathbf{s}) = \mu$ (desconocido)
- estacionariedad de segundo orden con covariograma $C(\mathbf{h})$, $\mathbf{h} \in \mathbb{R}^n$

Ahora, para garantizar la insesgadez hay que hacer la minimización con la restricción $\sum_i = 1^n \lambda_i = 1$. Se introduce un multiplicador de Lagrange $2m$, y se minimiza la función lagrangiana

$$L(\lambda_1, \dots, \lambda_n, m) = \sigma_k^2 = \boldsymbol{\lambda}' \boldsymbol{\Sigma} \boldsymbol{\lambda} - 2\boldsymbol{\lambda}' \mathbf{c} + \sigma^2 + 2m(\boldsymbol{\lambda}' \mathbf{1} - 1).$$

Para esto, se buscan los puntos en los que se anula el gradiente

$$\nabla(L) = 0 \iff 2(\boldsymbol{\Sigma} \boldsymbol{\lambda} - \mathbf{c} + m\mathbf{1}, \boldsymbol{\lambda}' \mathbf{1} - 1)' = (\mathbf{0}, 0)',$$

lo que conduce al sistema lineal

$$\begin{aligned}\boldsymbol{\Sigma} \boldsymbol{\lambda} + m\mathbf{1} &= \mathbf{c} \\ \mathbf{1}' \boldsymbol{\lambda} &= 1.\end{aligned}$$

En forma matricial:

$$\begin{pmatrix} \boldsymbol{\Sigma} & \mathbf{1} \\ \mathbf{1}' & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\lambda} \\ m \end{pmatrix} = \begin{pmatrix} \mathbf{c} \\ 1 \end{pmatrix}. \quad (\text{A.2})$$

Este sistema es compatible determinado, siempre que Σ sea definida positiva.

Es habitual restringirse a la hipótesis más débil de *estacionariedad intrínseca*, que involucra las diferencias del proceso, en lugar de los valores. Así, la condición $\sum_{i=1}^n \lambda_i = 1$ permite escribir:

$$\begin{aligned} \left(Z_0 - \sum_{i=1}^n \lambda_i Z_i \right)^2 &= Z_0^2 - 2Z_0 \sum_{i=1}^n \lambda_i Z_i + \left(\sum_{i=1}^n \lambda_i Z_i \right)^2 \pm \sum \lambda_i Z_i^2 \\ &= Z_0^2 \left(\sum \lambda_i \right) - 2Z_0 \sum_{i=1}^n \lambda_i Z_i + \sum \lambda_i Z_i^2 \\ &\quad - \frac{1}{2} \left(\sum \lambda_i Z_i^2 - 2 \sum \sum \lambda_i \lambda_j Z_i Z_j + \sum \lambda_j Z_j^2 \right) \\ &= \sum \lambda_i (Z_0 - Z_i)^2 - \frac{1}{2} \sum \sum \lambda_i \lambda_j (Z_i - Z_j)^2. \end{aligned}$$

Tomando esperanzas, y teniendo en cuenta que $2\gamma(\mathbf{s}_i - \mathbf{s}_j) \equiv \mathbb{V}[Z_i - Z_j] = \mathbb{E}[(Z_i - Z_j)^2]$ (por insesgadez),

$$\boxed{\sigma_k^2 = \mathbb{V}[\hat{Z} - Z_0] = 2\boldsymbol{\lambda}'\boldsymbol{\gamma} - \boldsymbol{\lambda}'\boldsymbol{\Gamma}\boldsymbol{\lambda}}, \quad (\text{A.3})$$

donde $\boldsymbol{\gamma} \equiv (\gamma(\mathbf{s}_0 - \mathbf{s}_1), \dots, \gamma(\mathbf{s}_0 - \mathbf{s}_n))'$ y $\boldsymbol{\Gamma}$ es la matriz $n \times n$ cuya componente (i, j) es $\gamma(\mathbf{s}_i - \mathbf{s}_j)$.

Notar que las Ecuaciones (A.1) y (A.3) están relacionadas y se pueden deducir recíprocamente utilizando la relación entre las funciones de covarianza y de semivariograma

$$\gamma(\mathbf{h}) = C(\mathbf{0}) - C(\mathbf{h}) = \sigma^2 - C(\mathbf{h}),$$

pues entonces,

$$\boldsymbol{\Gamma} = \left(\gamma(\mathbf{s}_i - \mathbf{s}_j) \right) = \left(C(\mathbf{0}) - C(\mathbf{s}_i - \mathbf{s}_j) \right) = \mathbf{1}'\sigma^2\mathbf{1} - \boldsymbol{\Sigma},$$

y

$$\boldsymbol{\gamma} = \sigma^2 \mathbf{1} - \mathbf{c}.$$

Cuando minimizamos la varianza de predicción expresada de esta forma, sujeta a la restricción $\sum \lambda_i = 1$, procedemos de forma análoga, minimizando el lagrangiano (notar que el coeficiente de Lagrange se introduce aquí con signo opuesto al caso anterior)

$$L(\lambda_1, \dots, \lambda_n, m) = \sigma_k^2 = 2\boldsymbol{\lambda}'\boldsymbol{\gamma} - \boldsymbol{\lambda}'\boldsymbol{\Gamma}\boldsymbol{\lambda} - 2m(\boldsymbol{\lambda}'\mathbf{1} - 1).$$

Anulando el gradiente:

$$\nabla(L) = 0 \iff 2(\boldsymbol{\gamma} - \boldsymbol{\Gamma}\boldsymbol{\lambda} - m\mathbf{1}, \boldsymbol{\lambda}'\mathbf{1} - 1)' = (\mathbf{0}, 0)',$$

lo que conduce al sistema lineal

$$\begin{aligned} \boldsymbol{\Gamma}\boldsymbol{\lambda} + m\mathbf{1} &= \boldsymbol{\gamma} \\ \mathbf{1}'\boldsymbol{\lambda} &= 1. \end{aligned}$$

O, en forma matricial:

$$\begin{pmatrix} \boldsymbol{\Gamma} & \mathbf{1} \\ \mathbf{1}' & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\lambda} \\ m \end{pmatrix} = \begin{pmatrix} \boldsymbol{\gamma} \\ 1 \end{pmatrix}. \quad (\text{A.4})$$

Notar que el Sistema (A.4) es formalmente idéntico al Sistema (A.2). De modo que la resolución se aborda a continuación con esta última notación, pero es igual de válida para ambos casos.

A.3 Inversión del sistema

Las operaciones son en bloques, y se hacen por fila, por tanto, las multiplicaciones son siempre por la izquierda.

$$\begin{aligned}
\left(\begin{array}{cc|cc} \Gamma & \mathbf{1} & \mathbf{1} & \mathbf{0} \\ \mathbf{1}' & 0 & \mathbf{0}' & 1 \end{array} \right) & \xrightarrow{\Gamma^{-1} \times F_1} \left(\begin{array}{cc|cc} \mathbf{1} & \Gamma^{-1}\mathbf{1} & \Gamma^{-1} & \mathbf{0} \\ \mathbf{1}' & 0 & \mathbf{0}' & 1 \end{array} \right) \\
& \xrightarrow{F_2 - \mathbf{1}'F_1} \left(\begin{array}{cc|cc} \mathbf{1} & \Gamma^{-1}\mathbf{1} & \Gamma^{-1} & \mathbf{0} \\ \mathbf{0}' & -\mathbf{1}'\Gamma^{-1}\mathbf{1} & -\mathbf{1}'\Gamma^{-1} & 1 \end{array} \right) \\
& \xrightarrow{F_1 + \frac{\Gamma^{-1}\mathbf{1}}{\mathbf{1}'\Gamma^{-1}\mathbf{1}}F_2} \left(\begin{array}{cc|cc} \mathbf{1} & \mathbf{0} & \Gamma^{-1} - \frac{\Gamma^{-1}\mathbf{1}\mathbf{1}'\Gamma^{-1}}{\mathbf{1}'\Gamma^{-1}\mathbf{1}} & \frac{\Gamma^{-1}\mathbf{1}}{\mathbf{1}'\Gamma^{-1}\mathbf{1}} \\ \mathbf{0}' & -\mathbf{1}'\Gamma^{-1}\mathbf{1} & -\mathbf{1}'\Gamma^{-1} & 1 \end{array} \right) \\
& \xrightarrow{\frac{-1}{\mathbf{1}'\Gamma^{-1}\mathbf{1}} \times F_2} \left(\begin{array}{cc|cc} \mathbf{1} & \mathbf{0} & \Gamma^{-1} - \frac{\Gamma^{-1}\mathbf{1}\mathbf{1}'\Gamma^{-1}}{\mathbf{1}'\Gamma^{-1}\mathbf{1}} & \frac{\Gamma^{-1}\mathbf{1}}{\mathbf{1}'\Gamma^{-1}\mathbf{1}} \\ \mathbf{0}' & 1 & \frac{\mathbf{1}'\Gamma^{-1}}{\mathbf{1}'\Gamma^{-1}\mathbf{1}} & \frac{-1}{\mathbf{1}'\Gamma^{-1}\mathbf{1}} \end{array} \right).
\end{aligned}$$

De modo que :

$$\begin{aligned}
\begin{pmatrix} \lambda \\ m \end{pmatrix} &= \begin{pmatrix} \Gamma^{-1} - \frac{\Gamma^{-1}\mathbf{1}\mathbf{1}'\Gamma^{-1}}{\mathbf{1}'\Gamma^{-1}\mathbf{1}} & \frac{\Gamma^{-1}\mathbf{1}}{\mathbf{1}'\Gamma^{-1}\mathbf{1}} \\ \frac{\mathbf{1}'\Gamma^{-1}}{\mathbf{1}'\Gamma^{-1}\mathbf{1}} & \frac{-1}{\mathbf{1}'\Gamma^{-1}\mathbf{1}} \end{pmatrix} \begin{pmatrix} \gamma \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} \Gamma^{-1}\gamma - \frac{\Gamma^{-1}\mathbf{1}\mathbf{1}'\Gamma^{-1}\gamma + \frac{\Gamma^{-1}\mathbf{1}}{\mathbf{1}'\Gamma^{-1}\mathbf{1}}}{\mathbf{1}'\Gamma^{-1}\mathbf{1}} \\ \frac{\mathbf{1}'\Gamma^{-1}\gamma - \frac{1}{\mathbf{1}'\Gamma^{-1}\mathbf{1}}}{\mathbf{1}'\Gamma^{-1}\mathbf{1}} \end{pmatrix} = \begin{pmatrix} \Gamma^{-1}(\gamma + \mathbf{1} \frac{1 - \mathbf{1}'\Gamma^{-1}\gamma}{\mathbf{1}'\Gamma^{-1}\mathbf{1}}) \\ \frac{\mathbf{1}'\Gamma^{-1}\gamma - 1}{\mathbf{1}'\Gamma^{-1}\mathbf{1}} \end{pmatrix}.
\end{aligned}$$

Finalmente, como Γ^{-1} es simétrica, y $\frac{\mathbf{1}'\Gamma^{-1}\gamma}{\mathbf{1}'\Gamma^{-1}\mathbf{1}}$ es un número, se puede escribir:

$$\boxed{\lambda' = (\gamma - \mu\mathbf{1})'\Gamma^{-1}, \quad m = -\frac{1 - \mathbf{1}'\Gamma^{-1}\gamma}{\mathbf{1}'\Gamma^{-1}\mathbf{1}}.}$$

Apéndice B

Glosario

Las definiciones relacionadas con la acústica se han extraído de [DOGV núm. 4394 \(2002\)](#), [BOE núm. 276/2003 \(2003\)](#), y [Parlamento y Consejo Europeos \(2002\)](#). Los términos relacionados con los Sistemas de Información Geográfica se han consultado en [Neteler y Mitasova \(2002\)](#), <http://www.gnu.org/philosophy/free-software-for-freedom.html> y <http://www.freegis.org>.

Capa (GIS) Una capa, dentro de una aplicación GIS, es una entidad que contiene datos geográficos (también llamados geodatos). Estos datos proveen una representación digital simplificada de una región geográfica, y se definen mediante una componente espacial (geométrica o gráfica) que describe su localización y otra componente de atributo que describe sus propiedades. Los geodatos pueden ser representados en base (principalmente) a 2

modelos de fuente de datos: modelo de datos *ráster* o modelo de datos *vectorial*.

Contaminación acústica Presencia en el ambiente de ruidos o vibraciones, cualquiera que sea el emisor acústico que los origine, que impliquen molestia, riesgo o daño para las personas, para el desarrollo de sus actividades o para los bienes de cualquier naturaleza, o que causen efectos significativos sobre el medio ambiente.

Datum El Datum es un conjunto de parámetros que permite definir la situación y orientación en el espacio de cualquier sistema de coordenadas. Además son parte del Datum las constantes físicas necesarias para definir de forma unívoca la posición de un punto en el espacio (masa terrestre, velocidad angular de rotación, etc.). Existen básicamente 2 formas de definir un Datum: mediante la definición física del sistema o bien mediante un conjunto de coordenadas que llevan implícito el modelo físico de referencia.

Decibelio Escala convenida habitualmente para medir la magnitud del sonido. El número de decibelios de un sonido equivale a 10 veces el valor del logaritmo decimal de la relación entre la energía asociada al sonido y una energía que se toma como referencia. Este valor también puede obtenerse de forma equivalente estableciendo la relación entre los cuadrados de las correspondientes presiones sonoras, en este caso el factor 10 veces deberá sustituirse por 20 veces ya que el logaritmo de un número al cuadrado es

igual al doble del logaritmo del citado número.

$$L_w = 10 \log_{10} \left(\frac{W}{W_{ref}} \right) \quad W = \text{potencia sonora}$$

$$L_r = 10 \log_{10} \left(\frac{I}{I_{ref}} \right) \quad I = \text{intensidad sonora}$$

$$L_p = 10 \log_{10} \left(\frac{P}{P_{ref}} \right)^2 = 20 \log_{10} \left(\frac{P}{P_{ref}} \right) \quad P = \text{presión sonora}$$

Indicador de ruido Una magnitud física para describir el ruido ambiental, que tiene una relación con un efecto nocivo.

Mapa acústico/de ruido La presentación de datos sobre una situación acústica existente o pronosticada en función de un *indicador de ruido*, en la que se indicará el rebasamiento de cualquier valor límite pertinente vigente, el número de personas afectadas en una zona específica o el número de viviendas expuestas a determinados valores de un indicador de ruido en una zona específica.

Proyección (cartográfica) Para transformar las coordenadas curvilíneas en coordenadas planas es necesario usar una proyección cartográfica. Proyecciones directas entre objetos esféricos y objetos planos no pueden ser acometidas sin distorsión. Muchas son las proyecciones desarrolladas en el intento de minimizar dichas distorsiones. Éstas pueden ser categorizadas en 3 grupos: *conformes* (que mantienen los valores angulares), *equidistantes* (manteniendo las distancias relativas) y *equivalentes* (manteniendo el área).

Ráster (modelo de datos) Un ráster es una matriz (o cuadrícula) regular de valores, y la forma de acceder a cada elemento del ráster (celda del ráster) es mediante sus números de fila y columna. Cada valor de la cuadrícula podrá referirse a una variable continua (elevación del terreno, temperatura, etc.) o a un conjunto de valores que representan atributos asociados a dicha celda (niveles digitales en una imagen satelital). La resolución de un ráster equivale a la longitud de la celda, tanto en sentido horizontal como vertical, y dicho parámetro controla el nivel de detalle de la representación espacial de los datos.

Ruido ambiental El sonido exterior no deseado o nocivo generado por las actividades humanas, incluido el ruido emitido por los medios de transporte, por el tráfico rodado, ferroviario y aéreo y por emplazamientos de actividades industriales como los descritos en el anexo I de la Directiva 96/61/CE del Consejo, de 24 de septiembre de 1996, relativa a la prevención y al control integrados de la contaminación.

Sistema de coordenadas La información que nutre un GIS son los geodatos (*ver Capa*). Dichos datos deben tener definida su localización geográfica haciendo uso de un sistema de coordenadas, bien sea curvilíneo (esférico, elipsoidal) o cartesiano (coordenadas planas). Cada sistema de coordenadas queda definido por un origen (primer meridiano, etc.), unos ejes coordenados (x e y; este y norte; etc.), y sus unidades de medida (grados, metros, pies, etc.).

Sistema de Información Geográfica Un sistema de información geográfica (GIS por sus siglas en inglés) es una aplicación informática mediante la cual se puede almacenar, visualizar, analizar y representar datos cuya principal característica es estar asociados a una localización geográfica.

Sistema de Referencia Es la conjunción de un *Sistema de Coordenadas* con un *Datum*.

Software Libre Richard M. Stallman fue el primero en definir el concepto de Software Libre a través de las siguientes cuatro *libertades*:

0. libertad de ejecución de la aplicación para cualquier propósito;
1. libertad de estudiar el funcionamiento de la aplicación, y adaptarla a cada necesidad;
2. libertad de redistribuir copias;
3. libertad para mejorar la aplicación, y distribuir dichas mejoras de forma que el conjunto de la comunidad se beneficie.

Las aplicaciones que cumplen estas cuatro libertades son llamadas Software Libre. Existen hoy en día varias licencias de software que garantizan dichas libertades. Una de las más conocidas y utilizadas es la GNU General Public Licence (<http://www.gnu.org/copyleft/gpl.html>).

Sonómetro Instrumento provisto de un micrófono amplificador detector de RMS, integrador-indicador de lectura y curvas de ponderación, que se utiliza para medición de niveles de presión sonora.

Vectorial (modelo de datos) Este modelo se basan en la representación *arco-nodo*. Cada arco se almacena como una secuencia de puntos definidos por sus coordenadas (x,y). Cada extremo del arco es llamado nodo. De esta forma es posible representar 3 figuras geométricas básicas: puntos, líneas y polígonos. A cada elemento representado gráficamente, le es asignado un identificador único a través del cual se le asocia información descriptiva (atributos). Frecuentemente el modelo vectorial incluye información *topológica* que refiere a las relaciones espaciales entre los elementos de una misma capa.

Apéndice C

Implementación del cálculo de distancias basadas en el coste

La implementación del procedimiento de cálculo de las matrices de distancia la hemos empaquetado en un script llamado `v.costdist.mat`, siguiendo la nomenclatura estándar de los *add-ons* de GRASS.

En este apéndice se incluye el código fuente y la referencia de uso del script, que también se pueden descargar de la URL <http://www.geeitema.org/doc/guenmap//docs/v.costdist.mat.zip>.

Referencia de uso

NAME

v.costdist.mat - Cost-based distances between points

KEYWORDS

SYNOPSIS

v.costdist.mat

v.costdist.mat help

```
v.costdist.mat [-k] from=filename
                  to=filename[,filename,...]
                  cost=filename
```

Flags:

-k

Use the 'Knight's move'. Forwarded to r.cost (see r.cost).

Parameters:

from=filename

Name of existing vector points map

to=filename[,filename,...]

Name(s) of existing vector points map(s)

where cost-distances to "from" points are to be stored

cost=filename

Name of existing raster file containing cost information

Código fuente

```

1  #!/bin/bash
2  #####
3  #
4  # MODULE:          v.costdist.mat
5  #                  version Sep 22, 2009
6  #
7  # AUTHOR:          Facundo Muñoz, Universitat de València, Spain
8  #
9  # PURPOSE:         Calculating cost-based distances between points
10 #                  see: A. López-Quílez & F. Muñoz (2009). ‘
11 #                  Geostatistical computing
12 #                  of acoustic maps in the presence of barriers’.
13 #                  Mathematical and Computer Modelling 50(5-6)
14 #                  :929-938.
15 #
16 # COPYRIGHT:       (c)
17 #                  This program is free software under the GNU General
18 #                  Public
19 #                  License (>=v2). Read the file COPYING that comes
20 #                  with GRASS
21 #                  for details.
22 #
23 #####
24 #
25 # Description:
26 # v.costdist.mat determines de cumulative cost of moving between
27 # each pair of points

```

```
23 # from two given maps over a cost surface
24 # e.g.
25 # - obtaining the "true" distance matrix of a set of noise
    # measures in a city
26 #     considering buildings as non-transparent areas for sound
27 #
28 # Notes:
29 # Results are stored as new columns in "to" map's attributes
    # tables
30 # Try and keep the number of input sites to under a few dozen, as
    # the
31 # process is very computationally expensive. You might consider
    # creating
32 # a few hundred MB RAM-disk for a temporary mapset to avoid
    # thrashing your
33 # hard drive (r.cost is heavy on disk IO).
34
35 #%Module
36 #% description: Cost-based distances between points
37 #%End
38
39 #%option
40 #% key: from
41 #% key_desc: filename
42 #% type: string
43 #% gisprompt: old,vector,vector
44 #% description: Name of existing vector points map
45 #% required : yes
46 #%end
```

```
47
48 #%option
49 #%key: to
50 #%key_desc: filename
51 #%type: string
52 #%gisprompt: old,vector,vector
53 #%multiple : yes
54 #%description: Name(s) of existing vector points map(s) where cost
    -distances to "from" points are to be stored
55 #%required : yes
56 #%end
57
58 #%option
59 #%key: cost
60 #%key_desc: filename
61 #%type: string
62 #%gisprompt: old,cell,raster
63 #%description: Name of existing raster file containing cost
    information
64 #%required : yes
65 #%end
66
67 #%flag
68 #%key: k
69 #%description: Use the 'Knight's move'. Forwarded to r.cost (see r
    .cost).
70 #%end
71
72 if [ -z $GISBASE ] ; then
```

```

73     echo "You have to be in GRASS to use this." 1>&2
74     exit 1
75 fi
76
77 if [ "$1" != "@ARGS_PARSED@" ] ; then
78     exec $GISBASE/bin/g.parser "$0" "$@"
79 fi
80
81     # NOTE: Checking existence of input files is carried out by
82     g.parser
83     # thus, unneeded here
84 echo
85 echo "v.costdist.mat — Cost-based Distance Matrix"
86 echo "Processing $GIS_OPT_FROM -> $GIS_OPT_TO    cost=$GIS_OPT_COST"
87 echo "_____ "
88
89 FROM_FILE="$GIS_OPT_FROM"
90 COST_FILE="$GIS_OPT_COST"
91
92 unset KNIGHT
93 if [ $GIS_FLAG_K -eq 1 ] ; then
94     KNIGHT="-k"
95 fi
96
97 TMP_FROM='g.tempfile pid=$$ '
98 if [ $? -ne 0 ] || [ -z "$TMP_FROM" ] ; then
99     echo "ERROR: unable to create temporary files" 1>&2
100    exit 1

```

```

101 fi
102
103 # crop out only points in region
104 v.in.region output=tmp_cdm_region_$$ > /dev/null 2>&1
105 v.select ainput="$FROM_FILE" binput=tmp_cdm_region_$$ \
106     atype=point btype=area output=tmp_cdm_from_$$ > /dev/null 2>&1
107
108 # save in the temporary file an ascii representation of 'from_file'
109 v.out.ascii tmp_cdm_from_$$ > "$TMP_FROM"
110
111 # find N: number of source points
112 N="cat "$TMP_FROM" | wc -l"
113
114 if [ "$N" -gt 52 ] ; then
115     echo "WARNING: Computation is expensive! Please consider fewer
116         points." 1>&2
117     echo "Press enter to continue.." 1>&2
118     read -t 30
119 fi
120
121 # checking existence of target columns in each of 'to_files'
122 # if they don't exist, creates; else verifies they are complete.
123 # hoy many files are we dealing with
124 TOFILENUM='echo $GIS_OPT_TO | tr ', '\n' | wc -l'
125
126 # prepare columns in attribute tables
127 nf=1
128 while [ "$nf" -le "$TOFILENUM" ]

```

```

129  do
130      TOFILE='echo "$GIS_OPT_TO" | cut -d',' -f$nf '
131
132      if [ -z "$TOFILE" ] ; then
133          echo 'ERROR: Empty file name! Did you leave a trailing comma
134              ?' 1>&2
135          exit 1
136      fi
137
138      # at the end we need as many columns as source points i.e.
139      #
140      # N
141      NCOL="db.columns "$TOFILE" | grep -c "cdm_[[:digit:]]" ""
142
143      if [ $NCOL -eq 0 ] ; then
144          echo ""
145          echo "*** Incorporating distance matrix to $TOFILE
146              ***"
147
148      # FIX: Now columns are created later, using the actual cat numbers
149      # of the from map
150      #
151      # nc=1
152      # while [ "$nc" -le "$N" ]
153      # do
154      #
155      #     # could be interesting to work in real or
156      #     # even integer precision?
157      #
158      #     # echo "ALTER TABLE $TOFILE ADD COLUMN
159      #     # cdm_$nc double precision" |
160      #
161      #     # db.execute
162      #     echo "ALTER TABLE $TOFILE ADD COLUMN
163      #     # cdm_$nc int" |

```

```

151 #                               db.execute
152 #                               nc='expr $nc + 1'
153 #                               done
154     elif [ $NCOL -eq $N ] ; then
155         echo "WARNING: distance matrix already exists in
156             $TOFILE and will be
157             overwritten" 1>&2
158     else
159         echo "ERROR: found inconsistent distance matrix in
160             $TOFILE" 1>&2
161         echo "$NCOL distance columns were found, out of $N.
162             Please do the cleaning"
163             1>&2
164         g.mremove vect=tmp_cdm_*_$$ > /dev/null 2>&1
165         rm -f "$TMP_FROM" > /dev/null 2>&1
166         exit 1
167     fi
168     nf='expr $nf + 1'
169 done
170 # gen cost maps for each source point
171 # echo "Generating cost maps.." 1>&2
172 # figure out which column cats are in (3 or 4 depending on if there
173 # is a z coordinate)
174 CATCOL='cat "$TMP_FROM" | head -1 | tr '|' '\n' | wc -1'
175 NUM=1
176 for POS in `cat "$TMP_FROM" ` ; do

```

```

176
177 #   echo "POS=[$POS]"
178   EASTING='echo "$POS" | cut -f1 -d"|"'
179   NORTHING='echo "$POS" | cut -f2 -d"|"'
180   CAT='echo "$POS" | cut -f"$CATCOL" -d"|"'
181
182   echo "site $NUM of $N  e=$EASTING  n=$NORTHING  cat=$CAT"
183
184   if [ -z "r.what input=$COST_FILE east_north=$EASTING,$NORTHING
      | grep -v "*" | ] ; then
185     echo "  Skipping, site lays outside of cost map." 1>&2
186     N='expr $N - 1'
187     continue
188   fi
189
190   r.cost $KNIGHT in=$COST_FILE out=cost_site.$CAT coordinate=
      $EASTING,$NORTHING 2>/dev/null>&1
191
192   # incorporating cost-distance values to columns in each of
      the "to_files"
193   nf=1
194   while [ "$nf" -le "$TOFILENUM" ]
195     do
196     TOFILE='echo "$GIS_OPT_TO" | cut -d',' -f$nf'
197     # if $TOFILE has column cdm_$CAT, we drop it out,
      and recreate it
198     if [ -n "db.columns table=$TOFILE | grep cdm_$CAT"
      " ] ; then

```

```

199         echo "ALTER TABLE $TOFILE DROP COLUMN cdm_$CAT" |
           db.execute
200     fi
201     echo "ALTER TABLE $TOFILE ADD COLUMN cdm_$CAT int"
           | db.execute
202     # now we populate it with the corresponding cost
           value
203     v.what.rast vector=$TOFILE raster=cost_site.$CAT
           column=cdm_$CAT 2>/dev/null>&1
204     #echo "writing column $CAT of file $TOFILE..."
205     #remember to clean out possible evident absurd
           values from v.what.rast
206     # (no problem with this, so far.)
207     nf='expr $nf + 1'
208     done
209
210     NUM=" `expr $NUM + 1` "
211 done
212
213
214 # Clean up
215
216 echo "Cleanup .." 1>&2
217 #g.mremove -f rast=cost_site.* | grep REMOVE
218 g.mremove -f vect=tmp_cdm_*_$S | grep REMOVE
219 rm -f "$IMP_FROM"
220
221 # done!

```

```
222 echo "Done! Wrote results to attribute tables of: ${GIS_OPT_TO}."
      1>&2
223 exit 0
```

Apéndice D

Adaptación del paquete `geoR`

Necesitamos adaptar los algoritmos geoestadísticos para poder utilizar las distancias basadas en el coste. En este apéndice se detallan todos los cambios realizados en el código fuente del paquete `geoR`.

El código fuente modificado puede descargarse de <http://www.geeitema.org/doc/guenmap//docs/modif.geoR.1.6-27.tar.gz>. Corresponde a la versión 1.6-27 del paquete `geoR`, y se puede instalar desde R como un paquete normal. También publicamos los datos del proyecto piloto (<http://www.geeitema.org/doc/guenmap//docs/noise.Rdata>) y un script en R demostrando el Kriging basado en el coste que utiliza el paquete modificado y los datos piloto (<http://www.geeitema.org/doc/guenmap/docs/short.demo.R>).

Tanto las funciones de `geoR` como de `gstat` realizan el cálculo de distancias entre los datos entre sí, entre los datos y las localizaciones de predicción, y

entre las localizaciones entre sí, a partir de sus coordenadas. Este cálculo está *integrado* dentro de las funciones que calculan el semivariograma, y que realizan la predicción. Más aún, `geoR` sí que calcula las matrices de distancias. En cambio `gstat`, para evitar trabajar con matrices grandes, evita el cálculo de la matriz completa, haciendo los cálculos a medida que va necesitando los datos.

En las descripciones de la estructura original de los algoritmos que siguen a continuación se han mantenido los nombres de las secciones de código en inglés para facilitar la identificación en el código fuente.

Compilación de paquetes en R

En la documentación de R se incluye el documento `R-admin` ([R Development Core Team, 2009](#)), donde se explica el conjunto de herramientas a instalar necesario para compilar librerías de R, y se describe el procedimiento. En el directorio `R_HOME/src/geoR` se guardan los fuentes del paquete, y para compilarlo basta con ejecutar `R CMD INSTALL src/geoR` desde línea de comandos. Eso genera la estructura de directorios que correspondiente en el directorio `R_HOME/library`, sobrescribiendo, si es necesario los binarios previamente instalados.

Identificación de las modificaciones necesarias

El paquete `geoR` tiene una serie de funciones en C, que quedan incluidas dentro de la librería `geoR.dll`. Allí hemos identificado las funciones que hacen cálculo de distancias:

loccoords distancias entre datos y localizaciones de predicción

distdiag distancias entre localizaciones de predicción

diffpairs distancias entre datos y diferencias entre sus valores

cordiag diagonal inferior de la matriz de distancias (o correlaciones) entre localizaciones

Además, en el directorio `R_HOME/src/geoR/R`, de la estructura fuente del paquete, hay una serie de funciones escritas en lenguaje R. `Cwrappers.R` tiene funciones que hacen de interfaz entre las funciones de R y las funciones de C. En particular éstas, que detectamos como más relevantes: `kriging.R` tiene las funciones de predicción como `krige.conv()`; `variogram.R` contiene las relacionadas con el semivariograma empírico y `likGRF.R` tiene las funciones de maximización de la verosimilitud, entre ellas `likfit()`.

Los cálculos de distancias se hacen con funciones de C por su rapidez. Sin embargo, a los efectos de las pruebas de este trabajo, se intenta introducir las modificaciones al código fuente al nivel más alto posible.

Modificaciones para el cálculo del semivariograma empírico

Estructura original

La función `variog()` está ubicada en el archivo fuente `variogram.R`. Esquemáticamente tiene la siguiente estructura, donde se destacan los puntos en los cuales se hace un cálculo de distancias:

1. Preliminaries
2. Angles
3. Basics
4. Box-Cox
5. Trend removal
6. Calculation of vector of distances

```
u <- as.vector(dist(as.matrix(coords)))
```
7. Calculation of vector of angles if needed
8. If default opts.: `options="bin"`, `bin.cloud="FALSE"`, `omnidirectional`
Calcula bins, a partir de los vectores `u` y `uvec` llama a la función de `C binit` (le **pasa las coordenadas** de las observaciones y **dentro de la función se calculan las distancias Euclídeas** para saber en qué bin se acumula el dato) para calcular los vectores con los resultados esenciales del semivariograma empírico:

- `vbin`: valor del estimador elegido para cada bin
- `cbin`: bin counts
- `sdbin`: bin sd

Atiende a la opción `pairs.min`. Monta el objeto `result` con: `u`, `v`, `n`, `sd`, `bins.lim`, `ind.bin`

9. General case:

calcula el vector `v` de estimadores, para cada par de observaciones, a partir de `u`. Filtros de distancia (`max.dist`) y ángulo (`direction`). Si `option='cloud'`, monta `result=(u,v)` y los ángulos si procede. Si `option='bin'`, llama a la función interna `.rfm.bin` que hace esencialmente lo mismo que la función de `C` `binit`: calcula bins, y los vectores `vbin`, `nbin`, `sdbin`, determinando la inclusión a partir del vector `u`. Luego atiende a las opciones `pairs.min`, `keep.NA` y `nugget.tolerance`, monta el objeto `result` con `u`, `v`, `n`, `sd`, `bins.lim`, `output.type='bin'`. Si `option='smooth'` aplica la función `ksmooth(u,v)` y devuelve el resultado en `result=(u,v)`.

10. Añade a `result` elementos comunes a todos los casos

11. END.

Cambios efectuados

Nuestro caso de uso, por ahora, se mueve dentro de la línea de las opciones por defecto. Por lo tanto bastaría con alterar la función de `C`. Sin embargo esto

implica pasar desde R a C todo el vector de distancias, y además resultaría en una modificación que sólo sirve a un caso particular. Preferimos hacer una modificación general, que permita utilizar distancias no euclídeas con todas las opciones disponibles en la función `variog()`.

En primer lugar se ha añadido un nuevo argumento `dists.mat` (sin opción por defecto), con el objeto de pasarle a `variog()` una matriz de distancias personalizada. Esto requiere además un bloque adicional de código destinado a detectar la presencia del argumento, y verificar que sea una matriz del tamaño adecuado.

En el punto 6, donde se efectúa el cálculo del vector `u` de distancias se ha introducido una bifurcación, de modo que si no está presente el argumento el cálculo se hace como antes, mientras que si está presente entonces el vector `u` se calcula a partir de la matriz proporcionada. Cabe apuntar que el vector de distancias sólo contiene la diagonal inferior de la matriz, contada por columnas. Es decir, que se asume implícitamente que $d(i, i) = 0$ y que $d(j, i) = d(i, j)$.

Finalmente, en el punto 8, que corresponde al caso con opciones por defecto, se introdujo otra bifurcación, otra vez dependiendo de la presencia del argumento nuevo. En el caso omiso se procede como antes, y en el caso de presencia del argumento **se evita la llamada a la función de C ‘binit’**, haciendo todos los cálculos dentro de R, en forma análoga a como se hacen en el caso general, de modo de aprovechar las distancias recogidas en el vector `u`.

Notar que el caso general permanece inalterado, porque efectivamente utiliza el vector `u` para determinar el bin donde corresponde incluir cada dato.

En resumen, son cuatro añadiduras en el código de `variogram.R`: un argumento nuevo, un bloque de código para controlarlo, y dos bifurcaciones que cambian el comportamiento normal en caso de presencia del argumento. Todas las líneas modificadas terminan con un comentario `# (GUENMAP)`, que puede servir de clave para buscar los cambios, y para volver al código original. Además, con estos cambios, la función `variog()` mantiene su funcionamiento inalterado cuando el argumento `dists.mat` no está presente.

Modificaciones para la estimación del modelo de semivariograma

Estructura original

La función `likfit()` está ubicada en el archivo fuente `likGRF.R`. Esquemáticamente tiene la siguiente estructura, donde se destacan los puntos en los cuales se hace un cálculo de distancias:

1. Check input
2. Setting coordinates, data and covariate matrices
3. Setting a factor to indicate different realizations
4. Initial values for parameters
5. Checking for multiple initial values for preliminar search of best initial value

Sólo en el caso de que haya más de un valor inicial de algún parámetro (lo cual no es nuestro caso). Calcula el vector de distancias entre datos utilizando la función `vecdist()` definida previamente. El resultado lo almacena en `.likGRF.dists.vec`, que lo define como variable global. A continuación llama a la función `loglik.GRF(compute.dists=FALSE)`, para cada una de las combinaciones posibles de los parámetros iniciales, a los efectos de quedarse con aquellos con mejor verosimilitud. Notar que no le pasa `.likGRF.dists.vec` como parámetro, sino que ésta variable es usada por `loglik.GRF()` por estar definida globalmente. Finalmente, borra `.likGRF.dists.vec`, del entorno global.

6. Box-Cox transformation for fixed lambda
7. Coordinates transformation for fixed anisotropy parameters

Si los parámetros de anisotropía (`psiR`, `psiA`) están fijos (`fix.psi.=TRUE`) (opción por defecto, y nuestro caso), entonces corrige la anisotropía si es necesario (`coords.aniso`) y calcula el vector de distancias con la función `vecdist()` (ya definida). Almacena el resultado en `.likGRF.dists.vec` en forma global. Este vector definido en este momento es el que afecta posteriormente al cálculo de la matriz de covarianzas. Además determina el rango de distancias en que se mueven los parámetros.

8. Constant term in the likelihood
9. Numerical minimization of the negative loglikelihood

Optimiza la función `.negloglik.GRF()`, usando las distancias almacenadas globalmente en `.likGRF.dists.vec`. En caso de que los parámetros de anisotropía

no estén fijos, entonces el vector de distancias no está calculado, y se hace dentro de `.negloglik.GRF()`

10. Transforming data according to the estimated lambda (Box-Cox) parameter

11. Transforming `coords` for estimated anisotropy (if the case)

Si los parámetros de anisotropía están fijos, elimina `.likGRF.dists.vec.` del entorno global, y no hay más nada que hacer. Si no, transforma coordenadas si hace falta, con los parámetros recién optimizados y define la función `rangevecdist()` que es como `vecdist()` pero se queda sólo con el rango de distancias, y la usa inmediatamente para calcular el rango de distancias en que se mueven los datos.

12. Computing estimated beta y τ^2 , σ^2 (if the case)

Utiliza la función `varcov.spatial()` (definida en `likGRF.R`) para calcular la matriz \mathbf{V} de varianzas-covarianzas, y una lista de resultados relacionados. A esta función se le pasan las coordenadas de los datos vía su argumento `coords`.

13. Likelihood results for the model without spatial correlation

Optimiza la función `.negloglik.boxcox()`, sin afectar para nada las distancias entre puntos.

14. Assigning names to the components of the mean vector beta

15. Computing residuals and predicted values (isolated components of the model)

Vuelve a llamar a `varcov.spatial()` para la predicción de la parte espacial, y también le pasa las coordenadas de los datos vía `coords`.

16. Some warning messages about particular possible results.
17. END.

Notas de interés:

- El vector de distancias (que habitualmente guarda en la variable `.likGRF.dists.vec`) contiene la mitad triangular inferior de la matriz de distancias, leída por columnas. Es decir: $\{d_{ij} \mid i = 1, \dots, n-1; j = i+1, \dots, n\}$.
- La función `vecdist(x) := as.vector(dist(x))`, donde `x` es una matriz de coordenadas de tamaño $n \times 2$, y devuelve un vector de longitud $\frac{n(n-1)}{2}$ con las distancias en el orden explicado antes.
- Todos los procesos anotados en el esquema los hace para cada una de las realizations, si es que hay más de una.

`Loglik.GRF()` está también ubicada en `likGRF.R`. Define nuevamente la función `vecdist()` y en caso de que exista anisotropía, (`psiR!=1 OR psiA!=0`), transforma coordenadas y recalcula el vector de distancias. En caso contrario, si el argumento `compute.dists=TRUE`, recalcula igual el vector de distancias (sin transformar coordenadas), y si no, no hace nada, y asume que ya existe un vector de distancias calculado y guardado en la variable global `.likGRF.dists.vec`. De hecho, desde `likfit()` se llama a esta función con `compute.dists=FALSE`, con lo cual, si no hay anisotropía, simplemente no hace nada. Finalmente, llama a la función `varcov.spatial()` pasándole el vector de distancias, para obtener la matriz de covarianzas.

`.negloglik.GRF()` está también ubicada en `likGRF.R`. En caso de que alguno de los parámetros de anisotropía no esté fijo, define nuevamente la función `vecdist()`, transforma coordenadas y recalcula el vector de distancias. En caso contrario no hace nada, y asume que ya existe un vector de distancias calculado y guardado en la variable global `.likGRF.dists.vec`. Finalmente, llama a la función `varcov.spatial()` pasándole el vector de distancias, para obtener la matriz de covarianzas.

`varcov.spatial()` está en `corcov.R`. Su objetivo es el cálculo de la matriz de varianzas-covarianzas de los datos. Se le pueden pasar las coordenadas, a partir de las cuales calcula el vector de distancias con `as.vector(dist(coords))`, o bien se le puede pasar directamente el vector de distancias. Ésta, a su vez, llama a la función `cov.spatial()` pasándole el vector de distancias. `cov.spatial()` simplemente evalúa la función de covarianza correspondiente en cada uno de los elementos del vector de distancias, por lo que ya no hay más cálculos de distancias dentro de ella, ni por debajo.

Cambios efectuados

En primer lugar, hemos querido centralizar el manejo de distancias, que estaba disperso en varios puntos de varias funciones. Concretamente, la función `vecdist()` se definía una vez dentro de `likfit()`, otra más dentro de `loglik.GRF()` y una vez más en `.negloglik.GRF()`. Así que se ha creado un nuevo archivo de funciones llamado `dists.R` en el cual colocamos sólo esa función, y eliminamos (con comentarios `##`) todas las definiciones previas.

Adicionalmente, en el punto 11 de `likfit()`, se define la función `rangevec-`
`dist()` que hacía lo mismo que `vecdist()` (sin llamarla) pero además calculaba
el rango. Así que se ha modificado para que llamara a `vecdist()`, en vez de
calcular el vector de distancias por su cuenta, de modo que use nuestra definición
de `vecdist()`.

Además se ha añadido un nuevo argumento `dists.mat` (sin opción por defecto)
con el objeto de pasarle a `likfit()` una matriz de distancias personalizada.
Esto requiere además un bloque adicional de código destinado a detectar la pre-
sencia del argumento, verificar que sea una matriz del tamaño adecuado, y si es
el caso guardarla en una variable global `.personal.definition.of.distances`
para su uso posterior. Esto a su vez exige eliminar la variable global del entorno
al finalizar el algoritmo, para que no afecte el uso posterior de `likfit()`, así que
hay un pequeño bloque al final a estos efectos.

En resumen, son cuatro cambios (3 eliminaciones, y una sustitución) en el
código de `likfGRF.R`, y tres añadiduras: un argumento nuevo, un bloque de
código para controlarlo y el bloque final de control de la variable global. Todas las
líneas modificadas terminan con un comentario `# (GUENMAP)`, que puede servir
de clave para buscar los cambios, y para volver al código original. Las líneas
originales comentadas terminan con `(GUENMAP - original uncommented)`.

En el archivo fuente `corcov.R` también hace falta un cambio: en los puntos
12 y 15 de `likfit()` se le pasan coordenadas a `varcov.spatial()`, para que
ella calcule internamente el vector de distancias, por lo que hace falta modificarla
también. Simplemente cambiamos una línea, para que hiciera el cálculo del vector

de distancias con la función `vecdist()`. Esta línea se puede encontrar por la clave (`GUENMAP`).

Finalmente, el archivo fuente `dists.R` contiene la función `vecdist()` modificada. En caso de que se detecte la existencia de la variable global `.personal-definition.of.distances` se obtiene el vector de distancias a partir de ella. En caso contrario se ejecuta la función original.

Modificaciones para la predicción

Estructura original

La función `krige.conv()` está ubicada en el archivo fuente `kriging.R`. Esquemáticamente tiene la siguiente estructura, donde se destacan los puntos en los cuales se hace un cálculo de distancias:

1. Preliminaries: verificaciones y asignaciones básicas
2. Reading input: asignación de valores de `krige.control()`
3. Reading output options: asignación de valores de `output.control()`
4. Checking input
5. Selecting locations inside border
6. Building the trend matrix
7. Anisotropy correction

8. Box-Cox transformation
9. Setting covariance parameters
10. Start Kriging calculations `Vcov <- varcov.spatial(corcov.R)`

Cálculo de la matriz de varianzas-covarianzas de las **observaciones**. Pasa las coordenadas y los parámetros del modelo de correlación. Dentro de `varcov.spatial()` se llama a la función `vecdist()` que **ya está adaptada** para utilizar la matriz personalizada en caso de que exista como variable global. Por lo que no hace falta modificar nada aquí, excepto asegurarse de que llegados a este punto, esté definida la variable global `.personal.definition.of.distances` y que contenga la matriz de distancias personalizadas entre las observaciones.

```
v0 <- loccoords(Cwrappers.R)
```

Cálculo de la matriz de **distancias entre observaciones y localizaciones**. Pasa las coordenadas de observaciones y localizaciones. Nota: `v0` queda de tamaño `nobs×nloc` (hay que trasponer la muestra)

```
v0 <- cov.spatial(corcov.R)
```

Cálculo de la matriz de **varianzas-covarianzas entre observaciones y localizaciones**. Pasa la matriz de distancias traspuesta `v0` y la función le aplica el modelo de correlación correspondiente. No hace falta tocar nada.

11. Sampling from the resulting distribution:

```
varcov <- varcov.spatial(corcov.R)
```

Cálculo de la matriz de varianzas-covarianzas de las **localizaciones**. Pasa las coordenadas de las localizaciones y los parámetros del modelo de correlación.

Nota: se necesitaría la **matriz de distancias de coste entre localizaciones**, cosa que no tenemos, y que sería muy costoso de conseguir, al menos con el esquema de funcionamiento actual. Por tanto, tal como están las cosas, **no podemos usar simulación**.

12. Backtransforming moments of the prediction distribution.
13. Setting classes and attributes.
14. END.

Cambios efectuados

Es necesario añadir dos nuevos argumentos `dd.dists.mat` y `dl.dists.mat` (sin opción por defecto, corresponden con `data-data`, y `data-locations`), con el objeto de pasarle a `krige.conv()` las matrices de distancias personalizadas. Esto requiere además un bloque adicional de código destinado a detectar la presencia de los argumentos, y verificar que sean matrices del tamaño adecuado. En caso de detectar la presencia de `dd.dists.mat`, se genera la variable global `.personal.definition.of.distances`.

En el punto 10, donde se genera `v0`, la matriz de distancias entre observaciones y localizaciones hemos insertado una ramificación. En caso de detectar la presencia de la matriz `dl.dists.mat` entonces evita la llamada a `loccoords()`, y simplemente asigna a `v0` dicha matriz traspuesta. En caso contrario se procede como antes, con la llamada a `loccoords`.

Finalmente, se requiere un pequeño bloque final para la limpieza de la variable global `.personal.definition.of.distances`.

Como siempre, todos los cambios están identificados con la etiqueta # (GUEN-MAP) para facilitar su ubicación, diagnóstico o modificación.

Limitaciones

No podemos usar simulación, en el esquema actual, porque es necesario disponer de la matriz de distancias `locations-locations` (27151×27151 !!!).

Hay que asegurarse de que todas las `locations` caen dentro de los límites de la región de `geodata`. Tal como está programado ahora, el objeto `geodata` no lleva información acerca de los límites de la región, o sea que no hay problema. Pero hay que tener en cuenta que `krige.conv()` elimina las `locations` fuera de la región, por lo tanto podría surgir un problema de incompatibilidad de dimensiones con la matriz de distancias no euclídeas.

Notas adicionales

`krige.conv()` no verifica que la matriz de varianzas-covarianzas de las observaciones sea definida positiva. Sobre el final, cuando calcula la varianza de predicción, entonces hace nulas todas las varianzas menores de $1e - 8$. En particular, si sale una negativa, termina siendo nula. Esto último no se comprueba, así que ocurre sin que el usuario se entere de que ha tenido una varianza negativa.

Una forma de analizar la salida consiste en editar la función `krige.conv()` con el comando `edit()` para introducir una llamada a `browser()` justo antes de que se efectúe esta *corrección* de varianzas. De esta manera se ha podido grabar el vector de varianzas en el archivo `var.anul.RData`, y analizar posteriormente su composición.

Quizá hubiera sido más acertado incluir los dos argumentos nuevos dentro de `krige.control()`, respetando la lógica de la función original.

Resumen de fuentes modificados

Fuente	Funciones modificadas
<code>dists.R</code> (nuevo)	<code>vecdist()</code> (nueva: cálculo selectivo de distancias)
<code>likfGRF.R</code>	<code>likfit()</code> , <code>loglik.GRF()</code> <code>.negloglik.GRF()</code>
<code>corcov.R</code>	<code>varcov.spatial()</code>
<code>variogram.R</code>	<code>variog()</code>
<code>kriging.R</code>	<code>krige.conv()</code>

Apéndice E

Geometría Riemanniana

E.1 Variedades Diferenciables

Aunque estas definiciones son estándar, se incluyen aquí por pertenecer al ámbito de la Geometría Diferencial y Riemanniana, y para establecer la notación. Algunas consecuencias se enuncian sin demostración, y nos ahorramos algo de rigor en las definiciones, para destacar los conceptos fundamentales. Para más detalles se pueden consultar las referencias [Adler y Taylor \(2007\)](#) o [Gudmundsson \(2009\)](#).

Una *Variedad Diferenciable* (Differentiable Manifold, en inglés) es una abstracción de objetos como curvas y superficies en \mathbb{R}^d . Intuitivamente puede verse como un conjunto *suave* de elementos con un sistema de coordenadas euclídeo definido localmente.

La definición rigurosa exige definir dicho sistema de coordenadas sobre un entorno de cada punto a través de un mapa o *carta* entre el entorno y un subconjunto de \mathbb{R}^n . Al conjunto de cartas que cubren el conjunto se le denomina *atlas*.

Definición E.1 (Atlas sobre un conjunto).

Sea un conjunto M cualquiera. Supongamos que podemos cubrir M con una familia de subconjuntos $U_\alpha \subset M$, $\cup_{\alpha \in I} U_\alpha = M$, donde I es un conjunto cualquiera de índices. En cada *trozo* U_α se puede establecer un sistema de coordenadas. Esto se hace a través de una correspondiente familia de aplicaciones biyectivas llamadas *cartas* $\phi_\alpha : U_\alpha \rightarrow A_\alpha \subset \mathbb{R}^n$, donde A_α es abierto, y $\alpha \in I$. La colección de cartas $\{\phi_\alpha : U_\alpha \rightarrow \mathbb{R}^n\}_{\alpha \in I}$ es un *atlas* de M .

Definición E.2 (Variedad Diferenciable).

Si el conjunto M es un espacio topológico Hausdorff localmente compacto tal que para todo punto $t \in M$ existe un abierto $U \subset M$ que lo contiene y un homeomorfismo $\phi : U \rightarrow \phi(U) \subset \mathbb{R}^n$, entonces se puede definir un atlas $\{\phi_\alpha : U_\alpha \rightarrow \mathbb{R}^n\}_{\alpha \in I}$ sobre M . Entonces M se denomina *Variedad Topológica*, y el número n es su *dimensión*.

Si además, para todo $\alpha, \beta \in I$ para los cuales $U_\alpha \cap U_\beta \neq \emptyset$ las *funciones de transición*

$$\phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha(U_\alpha \cap U_\beta) \rightarrow \phi_\beta(U_\alpha \cap U_\beta) \quad (\text{E.1})$$

son ellas y sus inversas k -veces diferenciables como funciones de subconjuntos de \mathbb{R}^n en \mathbb{R}^n (i.e., C^k -difeomorfismos), entonces el atlas es un C^k -atlas sobre M .

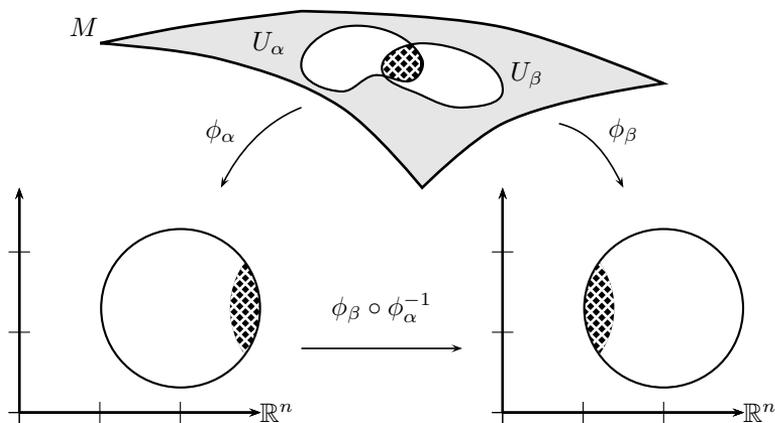


Figura E.1. Cartas sobre un conjunto M

Si además, el atlas no está contenido estrictamente en ningún otro atlas con cartas que satisfacen (E.1), entonces se dice que el atlas es *maximal*.

Una variedad topológica M de dimensión n , junto con un C^k -atlas maximal es una *Variedad C^k -diferenciable*.

La siguiente definición determina qué condiciones debe cumplir un subconjunto de una variedad diferenciable para ser él mismo una variedad diferenciable, heredando la estructura diferencial del espacio mayor.

Definición E.3 (Subvariedad y codimensión). Sean $m, n \in \mathbb{N}$ números naturales tales que $1 \leq m \leq n$, y sea N una C^r -variedad de dimensión n , con atlas maximal \mathcal{B} . Un subconjunto M de N es una *subvariedad* de N de dimensión n si para

todo punto $p \in N$ existe una carta $\phi_p : U_p \rightarrow \mathbb{R}^m \times \mathbb{R}^{n-m}$ que satisface

$$\phi_p(U_p \cup M) = \phi_p(U_p) \cup (\mathbb{R}^m \times 0),$$

donde U_p es un abierto que contiene al punto p . El número natural $n - m$ se llama la *codimensión* de M en N .

Si M es una subvariedad de N dotada de la topología restringida, y $\pi : \mathbb{R}^m \times \mathbb{R}^{n-m} \rightarrow \mathbb{R}^m$ es la proyección en el primer factor, entonces

$$\{\pi \circ \phi_p|_{U_p \cup M} : p \in M\}$$

es un C^r -atlas con el cual M es una C^r -variedad m -dimensional. La estructura diferenciable se llama *inducida* por la estructura diferenciable de N .

La noción de funciones continuas y diferenciables en una C^r -variedad es inmediata. Una función $f : M \rightarrow \mathbb{R}$ se dice que es de clase C^r si $f \circ \phi^{-1}$ lo es, en el sentido Euclídeo usual, para toda carta ϕ en el atlas.

E.2 Espacio Tangente

Para una variedad sumergida en \mathbb{R}^n , como una curva o una superficie, la idea de un vector tangente en un punto $p \in M$ es intuitiva y está formalizada en la Geometría Diferencial. Se trata, sencillamente, de un vector con origen en p contenido en el plano tangente a M en p . Dado tal vector v , es posible definir la

derivada de una función $f : M \rightarrow \mathbb{R}$ en la dirección de v . De este modo, cada vector v se corresponde con una derivada local para f .

En una variedad diferenciable abstracta M , la idea de vector tangente la construimos a partir del concepto de *curva diferenciable*. Una curva diferenciable en M que pasa por p es una aplicación diferenciable $\alpha : I \rightarrow M$, donde I es un intervalo en \mathbb{R} tal que $0 \in I$ y $\alpha(0) = p$.

Si U es un entorno abierto del punto $p \in M$, dada una función diferenciable $f : U \rightarrow \mathbb{R}$, se puede definir la derivada de f a lo largo de la curva α como

$$v(f) := (f \circ \alpha)'(0) \in \mathbb{R}. \quad (\text{E.2})$$

Se puede ver fácilmente que esta definición induce una relación de equivalencia entre las curvas que pasan por p . Cada una de dichas clases define un vector tangente determinado por acción sobre las funciones diferenciables. Esto es, el Funcional (E.2) que da, para cada función diferenciable el valor de su derivada en la *dirección* del vector.

El conjunto de vectores tangentes en un punto p forman un espacio vectorial llamado *Espacio Tangente* y denotado por $T_p M$.

Si consideramos la base canónica de \mathbb{R}^n , $\{e_i\}_{i=1}^n$, podemos definir las *curvas coordenadas* en M , por el punto p como $\tau_i(t) = \phi^{-1}(\phi(p) + te_i)$. Sus respectivas clases de equivalencia determinan sendos *vectores tangentes coordenados* $\{\frac{\partial}{\partial \phi^i} |_p\}$ que actúan sobre una función diferenciable como

$$\frac{\partial f}{\partial \phi^i} |_p = \frac{\partial (f \circ \phi^{-1})}{\partial x_i} |_{\phi(p)}.$$

Los vectores tangentes coordenados forman una base natural del espacio Tangente $T_p M$. Por tanto, la dimensión de dicho espacio es también $n = \dim M$.

E.3 Variedades Riemannianas

La estructura Riemanniana permite incorporar en las variedades las nociones de distancia, área o volumen.

Definición E.4 (Variedad Riemanniana). En una variedad diferenciable M , una *Métrica Riemanniana* es una familia de productos internos (en particular, definidos positivos)

$$g_p : T_p M \times T_p M \rightarrow \mathbb{R}, \quad p \in M.$$

Una variedad diferenciable dotada de una métrica Riemanniana se llama *Variedad Riemanniana*.

A pesar de su nombre, una métrica Riemanniana no es propiamente una métrica sobre M . Sin embargo, *induce* una métrica τ_g sobre M . Puesto que g determina la longitud de un vector tangente, podemos definir la longitud de una curva diferenciable $\alpha : [0, 1] \rightarrow M$ como

$$L(\alpha) = \int_0^1 \sqrt{g_{\alpha(t)}(\alpha'(t), \alpha'(t)')} dt$$

y definir la métrica τ_g como

$$\tau_g(p, q) = \inf\{L(\alpha) : \alpha \in D^1([0, 1]; M)_{(p, q)}\},$$

donde $D^1([0, 1]; M)_{(p, q)}$ es el conjunto de todas las aplicaciones diferenciables a trozos $\alpha : [0, 1] \rightarrow M$ con $\alpha(0) = p$, $\alpha(1) = q$.

Una curva diferenciable a trozos en M sobre la que se alcanza el ínfimo se llama una *geodésica* entre p y q . Las geodésicas no necesariamente son únicas.

Apéndice F

Sistema de coordenadas pseudo-hiperesféricas del espacio pseudo-Euclídeo

$$E(k, d-k)$$

Este sistema coordenado está adaptado para simplificar la ecuación del hiperboloide, que será del tipo $\rho = \text{cte}$. La coordenada ρ representa el *radio* del hiperboloide, es decir, la distancia pseudo-Euclídea al origen. El sistema tiene $k - 1$ ángulos *esféricos*, que tienen la misma interpretación que en el sistema de coordenadas esféricas. Además tiene $d - k$ *ángulos hiperbólicos* asociados a los

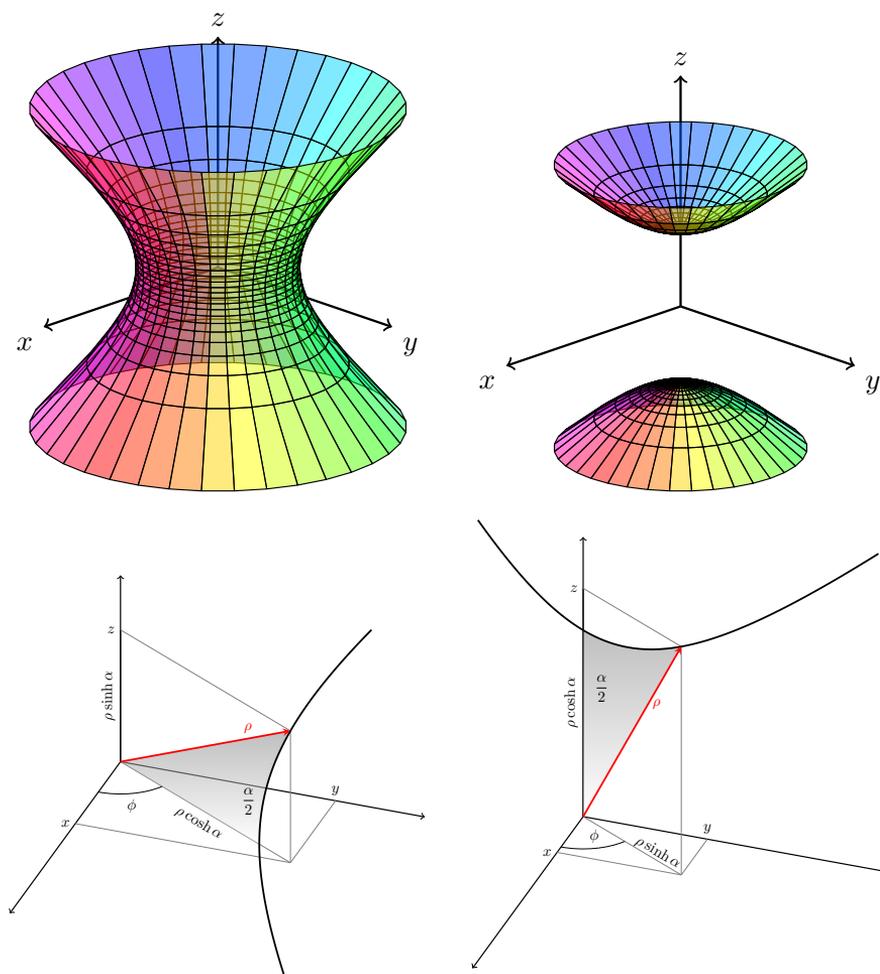


Figura F.1. Coordenadas cartesianas y pseudo-esféricas en $E_{(2,1)}$ para las regiones de cuadrado positivo (izquierda) y negativo (derecha).

ejes de *signo negativo*, que realmente no son ángulos, sino parámetros que varían en \mathbb{R} .

El espacio pseudo-Euclídeo tiene dos clases de *esferas*, o superficies isotrópicas. El hiperboloide formado por los puntos de cuadrado positivo, que en el espacio $E_{(2,1)}$ es el hiperboloide de una hoja de la Figura F.1 y el formado por los puntos de cuadrado negativo, que en $E_{(2,1)}$ es el hiperboloide de dos hojas.

En la Figura F.1 se representan estas coordenadas para el espacio $E_{(2,1)}$. El único ángulo esférico es el ángulo azimutal ϕ . El ángulo hiperbólico α representa el doble del área sombreada, y las proyecciones cartesianas de un punto de la hipérbola dependen del Seno y Coseno hiperbólicos de este parámetro.

La parametrización es distinta según se quiera representar puntos en la región de cuadrado positivo o negativo. No es posible recorrer todo el espacio pseudo-Euclídeo con una única parametrización de este tipo.

La generalización de este sistema de coordenadas a espacios de cualquier dimensión y signatura se definen en este apéndice. Luego se demuestran algunas relaciones elementales necesarias para calcular el jacobiano de la transformación, imprescindible a su vez para poder calcular integrales sobre la pseudo-hiperesfera (esto es, el hiperboloide) en el espacio pseudo-Euclídeo.

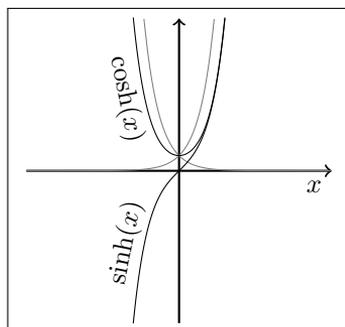
Será necesario utilizar dos parametrizaciones diferentes, una para cada región del espacio pseudo-Euclídeo, y comprobar todas las propiedades en cada caso. Muy en particular, los jacobianos de las respectivas transformaciones de coordenadas son diferentes.

Por completitud, presentamos la definición y algunos resultados elementales de las funciones hiperbólicas.

Definición F.1 (Seno y Coseno hiperbólicos).

$$\sinh(x) = \frac{e^x - e^{-x}}{2}$$

$$\cosh(x) = \frac{e^x + e^{-x}}{2}$$



La figura muestra las curvas y su relación con las funciones exponenciales (en gris claro).

Las siguientes relaciones se siguen inmediatamente de las definiciones:

$$\cosh^2(-x) - \sinh^2(x) = 1,$$

$$\sinh(-x) = -\sinh(x),$$

$$\frac{d}{dx} \sinh(x) = \cosh(x),$$

$$\cosh(-x) = \cosh(x),$$

$$\frac{d}{dx} \cosh(x) = \sinh(x).$$

F.1 Puntos de cuadrado positivo $(x, x) > 0$

Definición F.2 (Transformación de coordenadas pseudo-hiperesféricas a cartesianas).

Sean, $2 \leq k \leq d$. Dado $\rho > 0$, se define la siguiente transformación de coordenadas

$$T_{\rho}^{(+)} : \Phi = [0, 2\pi) \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]^{k-2} \times \mathbb{R}^{d-k} \rightarrow \mathbb{R}^d$$

$$T_{\rho}^{(+)}(\phi_2, \dots, \phi_k, \alpha_{k+1}, \dots, \alpha_d) = \mathbf{x}$$

cuyas componentes vienen dadas por las siguientes ecuaciones

$$\begin{aligned} \rho \cos \phi_2 \cos \phi_3 \cdots \cos \phi_{k-1} \cos \phi_k \cosh \alpha_{k+1} \cdots \cosh \alpha_{d-1} \cosh \alpha_d &= x_1 \\ \rho \sin \phi_2 \cos \phi_3 \cdots \cos \phi_{k-1} \cos \phi_k \cosh \alpha_{k+1} \cdots \cosh \alpha_{d-1} \cosh \alpha_d &= x_2 \\ \rho \sin \phi_3 \cdots \cos \phi_{k-1} \cos \phi_k \cosh \alpha_{k+1} \cdots \cosh \alpha_{d-1} \cosh \alpha_d &= x_3 \\ &\vdots \\ \rho \sin \phi_{k-1} \cos \phi_k \cosh \alpha_{k+1} \cdots \cosh \alpha_{d-1} \cosh \alpha_d &= x_{k-1} \\ \rho \sin \phi_k \cosh \alpha_{k+1} \cdots \cosh \alpha_{d-1} \cosh \alpha_d &= x_k \\ \rho \sinh \alpha_{k+1} \cdots \cosh \alpha_{d-1} \cosh \alpha_d &= x_{k+1} \\ &\vdots \\ \rho \sinh \alpha_{d-1} \cosh \alpha_d &= x_{d-1} \\ \rho \sinh \alpha_d &= x_d. \end{aligned} \tag{F.1}$$

Si $k = d$ se tienen coordenadas esféricas $(\rho, \phi_2, \dots, \phi_d)$. En el otro extremo, si $k = 1$, todas las variables son hiperbólicas $(\rho, \alpha_2, \dots, \alpha_d)$. En este caso, sin embargo, no hay ángulo azimutal que proporcione la simetría de revolución, y

las ecuaciones sólo parametrizan una mitad del hiperboloide. Específicamente, notar que x_1 es siempre mayor que ρ . Por tanto el caso $k = 1$ lo excluimos de la definición general, y lo tratamos en particular, al final de la sección.

La variable ρ representa la distancia (pseudo-Euclídea) al origen, y por tanto toma valores positivos. Las variables ϕ_j son ángulos en $[-\frac{\pi}{2}, \frac{\pi}{2}]$ excepto ϕ_2 que es un ángulo azimutal y varía entre 0 y 2π ; las variables α_j representan áreas con signo y pueden tomar cualquier valor real.

Alternativamente, se puede considerar ρ variando en todo \mathbb{R} y ϕ_2 variando en $[-\frac{\pi}{2}, \frac{\pi}{2}]$ igual que el resto de los ángulos esféricos. Esto permite simplificar algo la integración en el espacio completo, por lo que utilizaremos esta parametrización en algún caso. En cambio, se pierde la interpretación de ρ como el radio de la superficie isotrópica, que pasa a tener ecuación $|\rho| = \text{cte}$. El jacobiano sin embargo, es el mismo. Por brevedad, no repetiremos todas las demostraciones para este ligero cambio, ya que son fácilmente adaptables.

Para simplificar los cálculos que siguen, es conveniente adoptar la siguiente notación:

$$c_j = \begin{cases} 1, & j = 1 \\ \cos \phi_j, & 2 \leq j \leq k \\ \cosh \alpha_j, & k + 1 \leq j \leq d \end{cases} \quad s_j = \begin{cases} 1, & j = 1 \\ \sin \phi_j, & 2 \leq j \leq k \\ \sinh \alpha_j, & k + 1 \leq j \leq d \end{cases}, \quad (\text{F.2})$$

Si además adoptamos la convención de que si $m < n$ entonces $\prod_n^m \cdot = 1$, las ecuaciones del cambio de variables (F.1) se pueden expresar más resumidamente

como

$$x_i = \rho s_i \prod_{j=i+1}^d c_j, \quad 1 \leq i \leq d. \quad (\text{F.3})$$

En las siguientes demostraciones tendremos que discutir a menudo diversos casos. A los efectos de condensar argumentos que por otra parte son idénticos, vamos a utilizar la función indicatriz de un conjunto:

$$\mathbb{1}_{\{A\}}(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A. \end{cases}$$

Así, por ejemplo, el producto pseudo-Euclídeo lo podemos escribir como sigue

$$(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^k x_i y_i - \sum_{i=k+1}^d x_i y_i = \sum_{i=1}^d (-1)^{\mathbb{1}_{\{i>k\}}} x_i y_i.$$

Lema F.3. *Fijado un índice $2 \leq m \leq d$, se verifica la siguiente relación entre las variables cartesianas (x_m, \dots, x_d) y las pseudo-hiperesféricas (ρ, c_m, \dots, c_d)*

$$\rho^2 - \sum_{i=m}^d (-1)^{\mathbb{1}_{\{i>k\}}} x_i^2 = \rho^2 \prod_{j=m}^d c_j^2. \quad (\text{F.4})$$

Demostración. Para $m = d$,

$$\rho^2 - (-1)^{\mathbb{1}_{\{d>k\}}} \rho^2 s_d^2 = \rho^2 (1 - (-1)^{\mathbb{1}_{\{d>k\}}} s_d^2) = \rho^2 c_d^2,$$

ya que si $d > k$, entonces $s_d = \sinh \alpha_d$ y $1 + \sinh^2 \alpha_d = \cosh^2 \alpha_d$. En cambio, si $d = k$, $s_d = \sin \phi_k$, pero $1 - \sin^2 \phi_k = \cos^2 \phi_k$ y la relación se satisface en cualquier caso.

Supóngase que la relación es cierta para $3 \leq m \leq d$. Veámoslo para $m - 1$.

$$\begin{aligned} \rho^2 - \sum_{i=m-1}^d (-1)^{\mathbb{1}_{\{i>k\}}} x_i^2 &= \rho^2 - \sum_{i=m}^d (-1)^{\mathbb{1}_{\{i>k\}}} x_i^2 + (-1)^{\mathbb{1}_{\{m-1>k\}}} x_{m-1}^2 \\ &= \rho^2 \prod_{j=m}^d c_j^2 + (-1)^{\mathbb{1}_{\{m-1>k\}}} \rho^2 s_{m-1}^2 \prod_{j=m}^d c_j^2 \\ &= \rho^2 (1 + (-1)^{\mathbb{1}_{\{m-1>k\}}} s_{m-1}^2) \prod_{j=m}^d c_j^2 = \rho^2 \prod_{j=m-1}^d c_j^2. \end{aligned}$$

Por inducción, la relación (F.4) queda probada para $2 \leq m \leq d$. \square

El siguiente Lema prueba que el parámetro ρ efectivamente representa la distancia pseudo-Euclídea de un punto al origen.

Lema F.4 (Cuadrado de un vector).

Dado un vector $\mathbf{x} = (x_1, \dots, x_d) \in E_{(k, d-k)}$, su cuadrado tiene la siguiente expresión en coordenadas pseudo-hiperesféricas, que sólo depende de ρ .

$$(\mathbf{x}, \mathbf{x}) = \sum_{i=1}^k x_i^2 - \sum_{i=k+1}^d x_i^2 = \rho^2.$$

Demostración. Aplicando el Lema F.3 para $m = 2$ se tiene

$$\rho^2 - \sum_{i=2}^k x_i^2 + \sum_{i=k+1}^d x_i^2 = \rho^2 \prod_{j=2}^d c_j^2 = x_1^2.$$

De donde sale inmediatamente el resultado. \square

Así, la transformación $T_\rho^{(+)}$ lleva elementos de Φ a puntos de cuadrado positivo en \mathbb{R}^d sobre la superficie $d - 1$ dimensional de una pseudo-hiperesfera de radio ρ centrada en el origen

$$S_\rho = \{\mathbf{x} \in \mathbb{R}^d : x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_d^2 = \rho^2\}.$$

Veamos ahora que todo punto de cuadrado positivo puede ser expresado en términos de este sistema de coordenadas pseudo-hiperesféricas.

Descompóngase $S_\rho = H_\rho \cup F_\rho$ de forma disjunta, donde

$$H_\rho = \{\mathbf{x} \in S_\rho : x_1^2 + x_2^2 \neq 0\}$$

$$F_\rho = \{\mathbf{x} \in S_\rho : x_1^2 + x_2^2 = 0\}$$

y consideremos el dominio *abierto*

$$\begin{aligned} \Psi &= \{(\phi_2, \dots, \phi_k, \alpha_{k+1}, \dots, \alpha_d) \in \Phi : |\phi_j| \neq \frac{\pi}{2}, 3 \leq j \leq k\} \\ &= [0, 2\pi) \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)^{k-2} \times \mathbb{R}^{d-k}. \end{aligned}$$

Proposición F.5. Dado $\rho > 0$, la aplicación $T_\rho^{(+)} : \Psi \rightarrow H_\rho$ es biyectiva.

Demostración.

La aplicación está bien definida, ya que $T_\rho^{(+)}(\phi_2, \dots, \phi_k, \alpha_{k+1}, \dots, \alpha_d) \in S_\rho$ por el Lema F.4. Pero más específicamente, $T_\rho^{(+)}(\phi_2, \dots, \phi_k, \alpha_{k+1}, \dots, \alpha_d) \in H_\rho$:

$$x_1^2 + x_2^2 = \rho \prod_{j=3}^d c_j^2 > 0,$$

pues $\rho > 0$, $\cos \phi_j > 0 \forall \phi_j \in (-\frac{\pi}{2}, \frac{\pi}{2})$ y $\cosh \phi_j > 1 \forall \phi_j \in \mathbb{R}$.

Ahora, sea $(x_1, \dots, x_d) \in H_\rho$. Veamos que tiene una única preimagen en Ψ .

Puesto que la función $\sinh : \mathbb{R} \rightarrow \mathbb{R}$ es biyectiva, $\exists! \alpha_d \in \mathbb{R} : x_d = \rho \sinh \alpha_d$. Una vez determinado α_d , por el mismo argumento se encuentra un único α_{d-1} que satisface la relación $x_{d-1} = \rho \sinh \alpha_{d-1} \cosh \alpha_d$, y así sucesivamente hasta que se encuentran los únicos $\alpha_{k+1}, \dots, \alpha_d \in \mathbb{R}$ que satisfacen las correspondientes relaciones con las variables cartesianas x_{k+1}, \dots, x_d .

Ahora, como $(x_1, \dots, x_d) \in H_\rho \Rightarrow x_1^2 + \dots + x_k^2 = \rho^2 + x_{k+1}^2 + \dots + x_d^2$, con $x_1^2 + x_2^2 \neq 0$, entonces la siguiente desigualdad es estricta, y utilizando el Lema F.3,

$$x_k^2 \leq x_1^2 + x_2^2 + \dots + x_k^2 = \rho^2 + \sum_{i=k+1}^d x_i^2 = \rho^2 \prod_{j=k+1}^d \cosh^2 \alpha_j.$$

Por tanto se tiene que

$$-1 \leq \frac{x_k}{\rho \prod_{j=k+1}^d \cosh \alpha_j} \leq 1,$$

y entonces existe un único $\phi_k \in (-\frac{\pi}{2}, \frac{\pi}{2})$ tal que $\sin \phi_k = \frac{x_k}{\rho \prod_{j=k+1}^d \cosh \alpha_j}$, y que por tanto cumple la relación (F.3) para $i = k$.

Por inducción, si las coordenadas $\phi_{m+1}, \dots, \phi_k, \alpha_{k+1}, \dots, \alpha_d$, están determinadas para cierto $3 \leq m \leq k-1$, se argumenta de la misma manera,

$$x_m^2 \leq x_1^2 + x_2^2 + \dots + x_m^2 = \rho^2 - \sum_{i=m+1}^k x_i^2 + \sum_{i=k+1}^d x_i^2 = \rho^2 \prod_{j=m+1}^d c_j^2.$$

De modo que

$$-1 \leq \frac{x_m}{\rho \prod_{j=m+1}^d c_j} \leq 1,$$

y existe un único $\phi_m \in (-\frac{\pi}{2}, \frac{\pi}{2})$ tal que $\sin \phi_m = \frac{x_m}{\rho \prod_{j=m+1}^d c_j}$, y que por tanto cumple la relación (F.3) para $i = m$.

Por último,

$$x_1^2 + x_2^2 = \rho^2 - \sum_{i=3}^k x_i^2 + \sum_{i=k+1}^d x_i^2 = \rho^2 \prod_{j=3}^d c_j^2 = R^2 \geq 0.$$

Por tanto (x_1, x_2) son las coordenadas de un punto en una circunferencia de radio $R = \rho \prod_{j=3}^d c_j > 0$, y existe un único ángulo $\phi_2 \in [0, 2\pi)$ tal que $x_1 = R \cos \phi_2$ y $x_2 = R \sin \phi_2$.

Con todo esto, tenemos un único $(\phi_2, \dots, \phi_k, \alpha_{k+1}, \dots, \alpha_d) \in \Psi$ cuya imagen por $T_\rho^{(+)}$ sea (x_1, \dots, x_d) , lo que termina de probar la biyectividad de $T_\rho^{(+)}$. \square

Proposición F.6. Dado $\rho > 0$, y $(x_1, \dots, x_d) \in F_\rho$ existe un conjunto infinito (no numerable) de puntos en Φ cuya imagen por $T_\rho^{(+)}$ es (x_1, \dots, x_d) .

Demostración.

Por definición, $(x_1, \dots, x_d) \in F_\rho \Rightarrow x_1 = x_2 = 0$. Por tanto

$$x_1^2 + \dots + x_k^2 = x_3^2 + \dots + x_k^2 = \rho^2 + x_{k+1}^2 + \dots + x_d^2.$$

Por el mismo argumento utilizado en la Proposición F.5, existen únicos $\alpha_{k+1}, \dots, \alpha_d \in \mathbb{R}$ que verifican las ecuaciones (F.3) para $k+1 \leq i \leq d$.

Notar que si $k = 2$, $x_1^2 + x_2^2 = \rho^2 \cosh^2 \alpha_3 \cdots \cosh^2 \alpha_d \geq 0$. Por tanto, necesariamente $k \geq 3$ en este contexto. Ahora, contrariamente a la proposición anterior, la siguiente desigualdad no es estricta

$$x_k^2 \leq x_3^2 + \dots + x_k^2 = \rho^2 + \sum_{i=k+1}^d x_i^2 = \rho^2 \prod_{j=k+1}^d \cosh^2 \alpha_j.$$

Por tanto se tiene que

$$-1 \leq \frac{x_k}{\rho \prod_{j=k+1}^d \cosh \alpha_j} \leq 1.$$

Aún así, como la función Seno es biyectiva en $[-\frac{\pi}{2}, \frac{\pi}{2}]$, existe un único ϕ_k en dicho intervalo tal que $\sin \phi_k = \frac{x_k}{\rho \prod_{j=k+1}^d \cosh \alpha_j}$, y que por tanto cumple la relación (F.3) para $i = k$.

Ahora se distinguen dos casos.

$$\text{Caso 1. } \|x_k\| = \rho \prod_{j=k+1}^d \cosh \alpha_j.$$

Implica inmediatamente que $x_1 = \dots = x_{k-1} = 0$, y $\|\phi_k\| = \pi/2$. Por tanto $\cos \phi_k = 0$, y como este factor aparece en todas las ecuaciones del cambio de

variable, para $1 \leq i \leq k-1$, éstas se verifican trivialmente para cualquier elección de los ángulos $\phi_2, \dots, \phi_{k-1}$.

Notar que si $k = 3$ éste es el único caso posible.

Caso 2. $\|x_k\| < \rho \prod_{j=k+1}^d \cosh \alpha_j$.

Entonces existe al menos un $x_j \neq 0$ para cierto $3 \leq j \leq k-1$, y

$$\phi_k = \arcsin \left(\frac{x_k}{\rho \prod_{j=k+1}^d \cosh \alpha_j} \right) \in \left(-\frac{\pi}{2}, \frac{\pi}{2} \right),$$

con $\cos \phi_k > 0$.

Se repite el mismo razonamiento a partir de la desigualdad

$$x_{k-1}^2 \leq x_3^2 + \dots + x_{k-1}^2 = \rho^2 \prod_{j=k}^d c_j^2,$$

de donde se puede encontrar un único ángulo ϕ_{k-1} verificando la ecuación correspondiente, y se distinguen otra vez los dos casos según el valor de $\|x_{k-1}\|$.

Así sucesivamente hasta que, o bien se concluye que se tienen ángulos indeterminados, o bien se llega hasta

$$x_3^2 = \underbrace{x_1^2 + x_2^2}_{=0} + x_3^2 = \rho^2 - \sum_{i=4}^k x_i^2 + \sum_{i=k+1}^d x_i^2 = \rho^2 \prod_{j=4}^d c_j^2.$$

Puesto que por definición $x_3 = \rho \sin \phi_3 \prod_{j=4}^d c_j$, debe ser $\sin^2 \phi_3 = 1$, por tanto, $\phi_3 = \pm\pi/2$, y finalmente $\cos \phi_3 = 0$. De modo que las dos últimas ecuaciones se cumplen trivialmente, y ϕ_2 queda libre en $[0, 2\pi)$.

En cualquier caso, siempre se tiene al menos uno de los ángulos libres, lo que demuestra la proposición. \square

Cálculo del Jacobiano de la transformación $T_\rho^{(+)}$

De las ecuaciones (F.3) se pueden calcular fácilmente las derivadas respecto del parámetro radial,

$$\frac{dx_i}{d\rho} = s_i \prod_{j=i+1}^d c_j,$$

de los parámetros angulares,

$$\frac{dx_i}{d\phi_l} = \begin{cases} -\rho s_i \left(\prod_{j=i+1}^{l-1} c_j \right) s_l \left(\prod_{j=l+1}^d c_j \right) & i < l \\ \rho c_l \left(\prod_{j=l+1}^d c_j \right) & i = l \\ 0 & i > l \end{cases}$$

y de los ángulos hiperbólicos,

$$\frac{dx_i}{d\alpha_l} = \begin{cases} \rho s_i \left(\prod_{j=i+1}^{l-1} c_j \right) s_l \left(\prod_{j=l+1}^d c_j \right) & i < l \\ \rho c_l \left(\prod_{j=l+1}^d c_j \right) & i = l \\ 0 & i > l. \end{cases}$$

Notar que de cada una de las derivadas con respecto a los ángulos tanto esféricos como hiperbólicos se puede extraer un factor común $\rho \prod_{j=l+1}^d c_j$. En la matriz jacobiana, este factor es común a cada columna desde $l = 2$ hasta $l = d$, y por tanto puede extraerse fuera del determinante. De modo que ρ sale $d - 1$ veces, mientras que c_j lo hace $j - 2$ veces. Así, denotando genéricamente a las variables pseudo-hiperesféricas $\rho, \phi_2, \dots, \phi_k, \alpha_{k+1}, \dots, \alpha_d$ por $\omega_1, \dots, \omega_d$ respectivamente, y por f_j al factor común correspondiente, se tiene

$$J(T_\rho^{(+)}) = \rho^{d-1} \prod_{j=3}^d c_j^{j-2} \det \left(\frac{dx_i}{d\omega_j} / f_j \right).$$

Veremos a continuación que este último determinante tiene valor absoluto igual a 1, comprobando la siguiente igualdad matricial (Lema F.7):

$$\left(\frac{dx_i}{d\omega_j} / f_j \right)' \mathbf{S}_k \left(\frac{dx_i}{d\omega_j} / f_j \right) = \mathbf{S}_k,$$

donde \mathbf{S}_k es la matriz de signatura del espacio $E_{k,d-k}$ (ver Def. 6.5, p. 132).

Tomando determinantes, esto implica inmediatamente que

$$\det \left(\frac{dx_i}{d\omega_j} / f_j \right)' \det \mathbf{S}_k \det \left(\frac{dx_i}{d\omega_j} / f_j \right) = \pm 1,$$

y por tanto

$$\det \left(\frac{dx_i}{d\omega_j} / f_j \right) = \pm 1.$$

Y por tanto, el jacobiano de la transformación vale

$$\left| \det J(T_\rho^{(+)}) \right| = \left| \rho^{d-1} \prod_{j=3}^k (\cos \phi_j)^{j-2} \prod_{j=k+1}^d (\cosh \alpha_j)^{j-2} \right|.$$

Por armonía notacional podríamos escribir el primer productorio a partir del índice $j = 2$, gracias al exponente. Sin embargo, preferimos hacer explícito el hecho de que el ángulo ϕ_2 no interviene en el jacobiano.

Lema F.7. *Sea \mathbf{A} la matriz jacobiana del cambio de coordenadas pseudo-hiperesféricas para la región de cuadrado positivo en el espacio pseudo-Euclídeo $E_{k,d-k}$ dividida columna a columna por los correspondientes factores comunes*

$$\mathbf{A} = \left(\frac{dx_i}{d\omega_j} / f_j \right).$$

Entonces,

$$\mathbf{A}' \mathbf{S}_k \mathbf{A} = \mathbf{S}_k,$$

donde \mathbf{S}_k es la matriz de signatura del espacio.

Demostración. La prueba es una simple comprobación, pero resulta tediosa por la notación, y la diversidad de casos. Notar que el elemento (r, t) de la matriz denota el producto pseudo-Euclídeo de las columnas r y t de la matriz \mathbf{A} , que denotaremos por \mathbf{A}_r y \mathbf{A}_t respectivamente.

Comprobemos primero los elementos de la diagonal.

• **Caso 1:** $r = t = 1$. Este resulta el más sencillo, porque la primera columna de \mathbf{A} viene dada por las propias coordenadas cartesianas, divididas por el factor común ρ .

$$(\mathbf{A}_1, \mathbf{A}_1) = \frac{1}{\rho^2}(\mathbf{x}, \mathbf{x}) = \frac{1}{\rho^2} \left(\sum_{i=1}^k x_i^2 - \sum_{i=k+1}^d x_i^2 \right) = 1$$

por el Lema F.4.

• **Caso 2:** $2 \leq r = t \leq d$. Teniendo en cuenta que

$$\mathbf{A}_r(i) = \begin{cases} (-1)^{\mathbb{1}_{\{r \leq k\}}} s_i \prod_{j=i+1}^{r-1} c_j s_r, & i < r \\ c_r, & i = r \\ 0, & i > r \end{cases}$$

y utilizando inducción, se tiene

$$\begin{aligned} (\mathbf{A}_r, \mathbf{A}_r) &= s_r^2 \sum_{i=1}^{r-1} \underbrace{(-1)^{\mathbb{1}_{\{i > k\}}} \left(s_i \prod_{j=i+1}^{r-1} c_j \right)^2}_{=1} + (-1)^{\mathbb{1}_{\{r > k\}}} c_r^2 \\ &= (-1)^{\mathbb{1}_{\{r > k\}}} (c_r^2 + (-1)^{\mathbb{1}_{\{r > k\}}} s_r^2) = (-1)^{\mathbb{1}_{\{r > k\}}}. \end{aligned}$$

En efecto, para $r = 2$ se cumple la hipótesis

$$\sum_{i=1}^{r-1} (-1)^{\mathbb{1}_{\{i > k\}}} \left(s_i \prod_{j=i+1}^{r-1} c_j \right)^2 = s_1^2 = 1.$$

Y para $3 \leq r \leq d$ también se verifica. Utilizando inducción:

$$\begin{aligned} \sum_{i=1}^{r-1} (-1)^{\mathbb{1}_{\{i>k\}}} \left(s_i \prod_{j=i+1}^{r-1} c_j \right)^2 &= (-1)^{\mathbb{1}_{\{r-1>k\}}} s_{r-1}^2 + c_{r-1}^2 \underbrace{\sum_{i=1}^{r-2} (-1)^{\mathbb{1}_{\{i>k\}}} \left(s_i \prod_{j=i+1}^{r-2} c_j \right)^2}_{=1 \text{ (Hip. ind.)}} \\ &= c_{r-1}^2 + (-1)^{\mathbb{1}_{\{r-1>k\}}} s_{r-1}^2 = 1. \end{aligned}$$

Pasamos ahora a comprobar que las componentes fuera de la diagonal de este producto son nulas. Otra vez necesitamos distinguir diferentes casos.

• **Caso 3:** $r = 1 < t \leq d$.

$$\begin{aligned} (\mathbf{A}_1, \mathbf{A}_t) &= \sum_{i=1}^{t-1} (-1)^{\mathbb{1}_{\{i>k\}}} \left(s_i \prod_{j=i+1}^d c_j \right) \left((-1)^{\mathbb{1}_{\{t \leq k\}}} s_i \prod_{j=i+1}^{t-1} c_j s_t \right) \\ &\quad + (-1)^{\mathbb{1}_{\{t>k\}}} \left(s_t \prod_{j=t+1}^d c_j \right) \cdot c_t \\ &= (-1)^{\mathbb{1}_{\{t>k\}}} \left(s_t \prod_{j=t}^d c_j \right) \left(1 - \underbrace{\sum_{i=1}^{t-1} (-1)^{\mathbb{1}_{\{i>k\}}} \left(s_i \prod_{j=i+1}^{t-1} c_j \right)^2}_{=1} \right) = 0. \end{aligned}$$

Donde hemos utilizado nuevamente la inducción comprobada en el caso anterior.

- **Caso 4:** $2 < r < t \leq d$. Teniendo en cuenta que $\mathbf{A}_r(i) = 0$, $i > r$,

$$\begin{aligned}
(\mathbf{A}_r, \mathbf{A}_t) &= \sum_{i=1}^{r-1} (-1)^{\mathbb{1}_{\{i>k\}}} \left((-1)^{\mathbb{1}_{\{r \leq k\}}} s_i \prod_{j=i+1}^{r-1} c_j s_r \right) \left((-1)^{\mathbb{1}_{\{t \leq k\}}} s_i \prod_{j=i+1}^{t-1} c_j s_t \right) \\
&\quad + (-1)^{\mathbb{1}_{\{r>k\}}} c_r \left(s_r (-1)^{\mathbb{1}_{\{t \leq k\}}} \prod_{j=r+1}^{t-1} c_j s_t \right) \\
&= (-1)^{\mathbb{1}_{\{r \leq k < t\}}} \left(s_r s_t \prod_{j=r}^{t-1} c_j \right) \underbrace{\left[\sum_{i=1}^{r-1} (-1)^{\mathbb{1}_{\{i>k\}}} \left(s_i \prod_{j=i+1}^{r-1} c_j \right)^2 - 1 \right]}_{=1} = 0.
\end{aligned}$$

Otra vez por la inducción ya comprobada. \square

F.2 Caso $k = 1$

Como ya hemos visto, la parametrización dada por las ecuaciones (F.1) únicamente es válida cuando $k \geq 2$. En efecto, si fuera $k = 1$, no habría ningún ángulo esférico, y la parametrización solamente recorrería una mitad del hiperboloide, correspondiente a la rama positiva de la hipérbola base. En el caso $k = 1$, necesitamos considerar una parametrización alternativa.

Una posibilidad es utilizar las mismas ecuaciones, permitiendo que el parámetro ρ varíe en toda la recta real. En este escenario, la superficie isotrópica de radio fijo vendría dada por la ecuación $|\rho| = \text{constante}$. El jacobiano de la transformación es el mismo que para el caso general, ya que las ecuaciones son las mismas.

La otra parametrización, que resulta más natural es la siguiente:

$$\begin{aligned}
 x_1 &= \rho \cosh \alpha \\
 x_2 &= \rho \sinh \alpha \sin \phi_2 \\
 x_3 &= \rho \sinh \alpha \cos \phi_2 \sin \phi_3 \\
 &\vdots \\
 x_{d-1} &= \rho \sinh \alpha \cos \phi_2 \cos \phi_3 \cdots \cos \phi_{d-2} \sin \phi_{d-1} \\
 x_d &= \rho \sinh \alpha \cos \phi_2 \cos \phi_3 \cdots \cos \phi_{d-2} \cos \phi_{d-1}.
 \end{aligned}$$

Es inmediato comprobar por inducción que estos puntos están en la pseudo-hiperesfera en $E_{(1,d-1)}$:

$$x_1^2 - \sum_{i=2}^d x_i^2 = \rho^2 \left[\cosh^2 \alpha - \sinh^2 \alpha \underbrace{\left(\sum_{i=2}^d \left(\prod_{j=2}^{i-1} \cos^2 \phi_j \sin^2 \phi_i \right) \right)}_{=1} \right] = \rho^2,$$

donde hemos abusado de la notación considerando $\sin \phi_d = 1$.

Con esta parametrización, el parámetro α es el responsable de recorrer una rama de hipérbola, mientras que los parámetros angulares proporcionan la simetría de revolución alrededor de los ejes de signo negativo. El parámetro ρ debe ser negativo para recorrer la otra rama de la hipérbola. Así, $\rho \in \mathbb{R}$, y el radio de la superficie pseudo-hiperesférica es $|\rho|$. Los parámetros angulares pueden definirse en $[0, \pi)$, ya que por la forma de la hipérbola sólo hace falta *girar* media circunferencia en cada dirección.

F.3 Puntos de cuadrado negativo $(\mathbf{x}, \mathbf{x}) < 0$

Podemos aprovechar la elaboración anterior para definir la parametrización del espacio pseudo-Euclídeo en la región de puntos de cuadrado negativo. Basta notar que si invertimos el orden de los ejes, un punto de cuadrado positivo en $E_{(k, d-k)}$ se convierte en un punto de cuadrado negativo en $E_{(d-k, k)}$. En efecto, si $(\mathbf{x}, \mathbf{x}) = \sum_{i=1}^d (-1)^{\mathbb{1}_{\{i>k\}}} x_i^2 = \rho^2$, invertir el orden de los ejes equivale a considerar el punto \mathbf{y} tal que

$$y_i = x_{d-i+1}.$$

Y se tiene entonces, en $E_{(d-k, k)}$ con el cambio $j = d - i + 1$,

$$\begin{aligned} (\mathbf{y}, \mathbf{y}) &= \sum_{i=1}^d (-1)^{\mathbb{1}_{\{i>d-k\}}} y_i^2 = \sum_{i=1}^d (-1)^{\mathbb{1}_{\{i>d-k\}}} x_{d-i+1}^2 \\ &= \sum_{j=1}^d (-1)^{\mathbb{1}_{\{d-j+1>d-k\}}} x_j^2 = - \sum_{j=1}^d (-1)^{\mathbb{1}_{\{j>k\}}} x_j^2 = -\rho^2. \end{aligned}$$

Esto significa que podemos obtener una parametrización de la región de cuadrado negativo en $E_{(k, d-k)}$ invirtiendo el orden de las ecuaciones paramétricas (F.1) para el espacio $E_{(d-k, k)}$.

De esta manera *heredamos* todas las demostraciones hechas en la sección anterior. Esto es, sabemos que todos los puntos así parametrizados son de cuadrado negativo, y recíprocamente, que todo punto de cuadrado negativo

puede representarse en estas coordenadas de manera única (excepto los *polos* $x_{d-1}^2 + x_d^2 = 0$).

Las ecuaciones se definen consecuentemente de la siguiente manera.

Definición F.8 (Transformación de coordenadas pseudo-hiperesféricas a cartesianas).

Sean, $2 \leq d - k \leq d$. Dado $\rho > 0$, se define la siguiente transformación de coordenadas

$$T_\rho^{(-)} : \Phi = \mathbb{R}^k \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]^{d-k-1} \times [0, 2\pi) \rightarrow \mathbb{R}^d$$

$$T_\rho^{(-)}(\alpha_1, \dots, \alpha_k, \phi_{k+1}, \dots, \phi_{d-1}) = \mathbf{x}$$

cuyas componentes vienen dadas por las siguientes ecuaciones:

$$x_1 = \rho \sinh \alpha_1$$

$$x_2 = \rho \cosh \alpha_1 \sinh \alpha_2$$

$$\vdots$$

$$x_{k-1} = \rho \cosh \alpha_1 \cosh \alpha_2 \cdots \sinh \alpha_{k-1}$$

$$x_k = \rho \cosh \alpha_1 \cosh \alpha_2 \cdots \cosh \alpha_{k-1} \sinh \alpha_k$$

$$x_{k+1} = \rho \cosh \alpha_1 \cosh \alpha_2 \cdots \cosh \alpha_{k-1} \cosh \alpha_k \sin \phi_{k+1}$$

$$\vdots$$

$$x_{d-1} = \rho \cosh \alpha_1 \cosh \alpha_2 \cdots \cosh \alpha_{k-1} \cosh \alpha_k \cos \phi_{k+1} \cdots \cos \phi_{d-2} \sin \phi_{d-1}$$

$$x_d = \rho \cosh \alpha_1 \cosh \alpha_2 \cdots \cosh \alpha_{k-1} \cosh \alpha_k \cos \phi_{k+1} \cdots \cos \phi_{d-2} \cos \phi_{d-1}.$$

Notar que puesto que la parametrización anterior requería $k \geq 2$, la presente parametrización requiere $d \geq k + 2$, pues está basada en $E_{(d-k,k)}$.

Hemos invertido también el orden de la numeración de las variables pseudohiperesféricas según el cambio $j = d - i + 1$ por armonía notacional. Así, una versión condensada de las ecuaciones puede darse como

$$x_i = \rho s_i \prod_{j=1}^{i-1} c_j, \quad 1 \leq i \leq d.$$

utilizando la convención

$$c_j = \begin{cases} \cosh \alpha_j, & 1 \leq j \leq k \\ \cos \phi_j, & k + 1 \leq j \leq d - 1 \\ 1, & j = d \end{cases} \quad s_j = \begin{cases} \sinh \alpha_j, & 1 \leq j \leq k \\ \sin \phi_j, & k + 1 \leq j \leq d - 1 \\ 1, & j = d \end{cases}$$

Finalmente, también tenemos el cálculo del jacobiano, que tiene la misma forma que el anterior, teniendo en cuenta el cambio en los nombres de las variables:

$$\left| \det J(T_\rho^{(-)}) \right| = \left| \rho^{d-1} \prod_{j=1}^k (\cosh \alpha_j)^{d-j-1} \prod_{j=k+1}^{d-2} (\cos \phi_j)^{d-j-1} \right|.$$

Por último, el caso $d - k = 1$ es el análogo al caso $k = 1$. Notar que en la parametrización de la región de cuadrado negativo, en la Figura F.1 hemos

utilizado la parametrización alternativa

$$\begin{cases} x_1 = \rho \sinh \alpha \cos \phi \\ x_2 = \rho \sinh \alpha \sin \phi \\ x_3 = \rho \cosh \alpha, \end{cases}$$

que tiene Jacobiano

$$\left| \det J(T_\rho^{(-)}) \right| = \rho^2 |\sinh \alpha|.$$

F.4 Rotaciones del espacio pseudo-Euclídeo

Una *rotación* en un espacio de Hilbert (i.e., un espacio vectorial con producto interno completo) es una transformación lineal isométrica (i.e., que conserva el producto interno) cuya matriz tiene determinante 1. Forma parte de la familia de las transformaciones isométricas, que se caracterizan por tener matrices ortogonales de determinante ± 1 . Las otras transformaciones ortogonales son las simetrías.

En el espacio Euclídeo una rotación adecuada nos permite orientar un vector \mathbf{x} en cualquier dirección prefijada. En particular es posible orientarlo a lo largo de la dimensión j , de modo que $\mathbf{x}' = (x_i)$, $i = 1, \dots, d$ con $x_i = \delta_{ij} \|\mathbf{x}\|$.

El concepto se puede generalizar al espacio pseudo-Euclídeo, que aunque no tiene un producto interno propiamente, tiene un pseudo-producto interno.

Para simplificar la exposición, trabajemos en el espacio $E_{(2,1)}$, y supongamos que tenemos un vector \mathbf{x} , de cuadrado positivo, con coordenadas

$$\begin{cases} x = \rho \cos \phi \cosh \alpha \\ y = \rho \sin \phi \cosh \alpha \\ z = \rho \sinh \alpha, \end{cases}$$

que queremos orientar en la dirección del eje y .

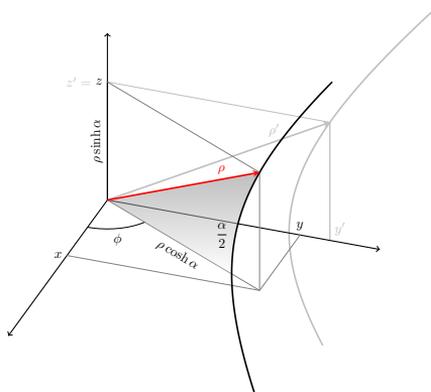


Figura F.2. Rotación de un vector de cuadrado positivo en el $E_{(2,1)}$.

Podemos descomponer la transformación en dos etapas (Fig. F.2). En primer lugar, una rotación azimutal en sentido antihorario de magnitud $\frac{\pi}{2} - \phi$ llevará la rama de la hipérbola sobre la que se encuentra el punto al plano (y, z) . Esto es

una transformación de matriz

$$\begin{bmatrix} \cos(\frac{\pi}{2} - \phi) & -\sin(\frac{\pi}{2} - \phi) & 0 \\ \sin(\frac{\pi}{2} - \phi) & \cos(\frac{\pi}{2} - \phi) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Las coordenadas del vector transformado son:

$$\begin{cases} x' = 0 \\ y' = \rho \cosh \alpha \\ z' = \rho \sinh \alpha. \end{cases}$$

Naturalmente, esta transformación ha preservado el pseudo-producto interno:

$$x'^2 + y'^2 - z'^2 = \rho^2(\cosh^2 \alpha - \sinh^2 \alpha) = \rho^2.$$

Ahora, se puede llevar a cabo una *rotación* a lo largo de la hipérbola, en dirección hacia el centro, de magnitud α , cuya matriz tiene la siguiente forma:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cosh(\alpha) & -\sinh(\alpha) \\ 0 & -\sinh(\alpha) & \cosh(\alpha) \end{bmatrix},$$

siendo las coordenadas transformadas:

$$\begin{cases} x'' = 0 \\ y'' = \rho \\ z'' = 0. \end{cases}$$

Al contrario que en el caso Euclídeo, sólo podemos orientar vectores de cuadrado positivo a lo largo de los ejes *de signo positivo*. Si no fuera así, estaríamos cambiando el valor del cuadrado del vector, y la transformación no conservaría esta magnitud. Podría plantearse una transformación que intercambiase los ejes, y la transformación seguiría siendo ortogonal. Pero en ese caso cambiaríamos del espacio $E_{(2,1)}$ al $E_{(1,2)}$, y esto no es deseable.

En el caso de los vectores de cuadrado negativo se procede de la misma manera, con la salvedad de que las dos ramas de las hipérbolas generatrices están *desconectadas*. En este caso, la parametrización contempla que $\rho \in \mathbb{R}$, de modo que cuando es positivo la parametrización recorre la rama $z > 0$ y cuando es negativo la rama $z < 0$. Así el vector transformado $z'' = \rho$ será colineal con el eje z siempre que el original estuviera en la rama *positiva*. En caso contrario, quedará orientado en sentido inverso.

En otras palabras, no basta con una rotación para orientar un vector de cuadrado negativo en la dirección (y sentido) del eje z . Se hace necesario admitir otra transformación ortogonal: una simetría. Aunque no es una rotación, también conserva las distancias cuadráticas con signo. A los efectos de integrar sobre todo el espacio el resultado es invariante ante este tipo de transformaciones.

Bibliografía

Abramowitz, M. y Stegun, I. A. (1972). *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*. Dover books on mathematics. Dover, New York. [16](#), [26](#), [30](#), [34](#), [154](#), [157](#), [158](#)

Adler, R. J. (1981). *The Geometry of Random Fields (Wiley Series in Probability and Statistics)*. John Wiley & Sons Inc. [105](#)

Adler, R. J. y Taylor, J. E. (2007). *Random Fields and Geometry (Springer Monographs in Mathematics)*. Springer, 1.^a edición. [109](#), [116](#), [255](#)

Awaida, A. y Westervelt, J. (2006). *r.cost: cumulative cost computation for GRASS GIS*. URL http://grass.osgeo.org/grass63/manuals/html63_user/r.cost.html. [72](#)

Banerjee, S., Carlin, B. P., y Gelfand, A. E. (2004). *Hierarchical Modeling and Analysis for Spatial Data*. Monographs on Statistics and Applied Probability. Chapman and Hall/CRC, 1.^a edición. [19](#)

- Barnard, J., McCulloch, R., y Meng, X.-L. (2000).** Modeling covariance matrices in terms of standard deviations and correlations, with application to shrinkage. *Statistica Sinica*, 10(4):1281–1311. [180](#)
- Bayarri, M. J., Berger, J. O., Cafeo, J., Garcia-Donato, G., Liu, F., Palomo, J., Parthasarathy, R. J., Paulo, R., Sacks, J., y Walsh, D. (2007).** Computer model validation with functional output. *The Annals of Statistics*, 35(5):1874–1906. [93](#)
- Berg, C., Christensen, J. P. R., y Ressel, P. (1984).** *Harmonic Analysis on Semigroups: Theory of Positive Definite and Related Functions*, volume 100 of *Graduate Texts in Mathematics*. Springer, 1.^a edición. [115](#), [116](#), [119](#), [205](#)
- Bhat, K. S., Haran, M., y Goes, M. (2010).** *Computer Model Calibration with Multivariate Spatial Output: A Case Study*, páginas 168–184. Springer. [93](#), [205](#)
- Bhatia, R. (2007).** *Positive Definite Matrices (Princeton Series in Applied Mathematics)*. Princeton University Press. [109](#)
- Bivand, R. (2008).** *spgrass6: Interface between GRASS 6 geographical information system and R*. URL <http://cran.r-project.org/package=spgrass6>. [70](#)
- Bochner, S. (1941).** Hilbert distances and positive definite functions. *The Annals of Mathematics*, 42(3):647–656. [116](#), [117](#), [119](#), [206](#)

- Bochner, S. (1955).** *Harmonic analysis and the theory of probability*. University of California Press, Berkeley and Los Angeles, CA. 15, 32
- Bochner, S. (1959).** *Lectures on Fourier Integrals*. Princeton University Press. 21, 26, 32
- BOE núm. 276/2003 (2003).** Ley 37/2003, de 17 de noviembre, del ruido. URL <http://www.boe.es/boe/dias/2003/11/18/index.php>. 219
- Boisvert, J. B., Manchuk, J. G., y Deutsch, C. V. (2009).** Kriging in the presence of locally varying anisotropy using Non-Euclidean distances. *Mathematical Geosciences*, 41(5):585–601. 208
- Borg, I. y Groenen, P. J. F. (2005).** *Modern Multidimensional Scaling: Theory and Applications (Springer Series in Statistics)*. Springer, 2.^a edición. 122
- Chilès, J.-P. y Delfiner, P. (1999).** *Geostatistics: Modeling Spatial Uncertainty*. Wiley Series in Probability and Statistics. Wiley-Interscience, 1.^a edición. 7
- Chiu, T. Y. M., Leonard, T., y Tsui, K. W. (1996).** The Matrix-Logarithmic covariance model. *Journal of the American Statistical Association*, 91(433):198–210. 179
- Christakos, G. (2000).** *Modern Spatiotemporal Geostatistics (International Association for Mathematical Geology: Studies in Mathematical Geology)*. Oxford University Press, USA. 114

- Christakos, G. y Papanicolaou, V. (2000).** Norm-dependent covariance permissibility of weakly homogeneous spatial random fields and its consequences in spatial statistics. *Stochastic Environmental Research and Risk Assessment*, 14(6):471–478. [114](#)
- Christakos, G. y Serre, M. L. (2000).** Spatiotemporal analysis of environmental exposure–health effect associations. *Journal of Exposure Science and Environmental Epidemiology*, 10(2):168–187. [208](#)
- Christakos, G., Hristopulos, D. T., y Bogaert, P. (2000).** On the physical geometry concept at the basis of space/time geostatistical hydrology. *Advances in Water Resources*, 23(8):799–810. [114](#)
- Cressie, N. (1993).** *Statistics for Spatial Data*. Wiley Series in Probability and Statistics. Wiley-Interscience, New York. [7](#), [17](#), [39](#)
- Cressie, N. y Hawkins, D. M. (1980).** Robust estimation of the variogram: I. *Mathematical Geology*, 12(2):115–125. [13](#)
- Cressie, N. y Zimmerman, D. L. (1992).** On the stability of the geostatistical method. *Mathematical Geology*, 24(1):45–59. [37](#)
- Cuadras, C. M., Fortiana, J., y Oliva, F. (1998).** *Anàlisi de dades multivariants*. Universitat de Barcelona, 3.^a edició. [122](#)
- Curriero, F. C. (1996).** *The use of non-Euclidean distances in geostatistics*. PhD thesis, Department of Statistics, Kansas State University, 213 p. [53](#)

- Curriero, F. C. (2006).** On the use of Non-Euclidean distance measures in geostatistics. *Mathematical Geology*, 38(8):907–926. [53](#), [58](#), [60](#), [114](#), [130](#), [143](#)
- Curriero, F. C., Hohn, M. E., Liebhold, A. M., y Lele, S. R. (2002).** A statistical evaluation of non-ergodic variogram estimators. *Environmental and Ecological Statistics*, 9(1):89–110. [13](#)
- Daniels, M. J. y Pourahmadi, M. (2009).** Modeling covariance matrices via partial autocorrelations. *Journal of multivariate analysis*, 100(10):2352–2363. [179](#)
- Dattorro, J. (2010).** *Convex Optimization & Euclidean Distance Geometry*. Meboo Publishing USA, Palo Alto, California. [207](#)
- Diggle, P. J. y Ribeiro, P. J. (2007).** *Model-based Geostatistics (Springer Series in Statistics)*. Springer Series in Statistics. Springer, New York, 1.^a edición. [17](#), [38](#), [39](#)
- DOGV núm. 4394 (2002).** Ley 7/2002, de 3 de diciembre, de la Generalitat Valenciana, de Protección contra la Contaminación Acústica [2002/13497]. URL <http://www.boe.es/boe/dias/2003/11/18/index.php>. [219](#)
- Federer, H. (1978).** Colloquium lectures on geometric measure theory. *Bulletin of the American Mathematical Society*, 84(3):291–339. [119](#), [206](#)
- Gihman, I. I. y Skorohod, A. V. (1974).** *The Theory of Stochastic Processes*. Springer-Verlag. [8](#), [20](#), [21](#), [26](#)

- Gneiting, T., Sasvári, Z., y Schlather, M. (2001).** Analogies and correspondences between variograms and covariance functions. *Advances in Applied Probability*, 33(3):617–630. [12](#)
- Goldfarb, L. (1985).** A new approach to pattern recognition. En Kanal, L. N. y Rosenfeld, A., editores, *Progress in Pattern Recognition*, volume 2, páginas 241–402. Elsevier Science Publishers BV. [137](#)
- Gower, J. C. (1966).** Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, 53(3/4):325–338. [122](#)
- Gower, J. C. (1968).** Adding a point to vector diagrams in multivariate analysis. *Biometrika*, 55(3):582–585. [138](#)
- Gower, J. C. (1982).** Euclidean distance geometry. *The Mathematical Scientist*, 7:1–14. [126](#)
- Gower, J. C. (1985).** Properties of euclidean and non-Euclidean distance matrices. *Linear Algebra and its Applications*, 67:81–97. [125](#)
- Gower, J. C. y Legendre, P. (1986).** Metric and euclidean properties of dissimilarity coefficients. *Journal of Classification*, 3(1):5–48–48. [124](#)
- Gradshteyn, I. S. y Ryzhik, I. M. (2007).** *Table of Integrals, Series, and Products*. Academic Press, 7.^a edición. [30](#), [33](#), [36](#), [154](#), [155](#), [163](#), [164](#), [165](#), [169](#), [173](#)

- GRASS Development Team (2008).** *Geographic Resources Analysis Support System (GRASS GIS) Software*. Open Source Geospatial Foundation, USA. URL <http://grass.osgeo.org>. 70
- Gudmundsson, S. (2009).** An introduction to Riemannian Geometry. Lecture Notes, Lund University. URL <http://www.matematik.lu.se/matematiklu/personal/sigma/>. 255
- Guttorp, P. y Gneiting, T. (2006).** Studies in the history of probability and statistics xlix: On the Matérn correlation family. *Biometrika*, 93(4):989–995. 17, 188
- gvSIG Development Team (2008).** *gvSIG Desktop, version 1.1.1, build number 1015*. Conselleria de Infraestructuras y Transporte, Valencia, España. URL <http://www.gvsig.org/>. 70
- Handcock, M. S. y Stein, M. L. (1993).** A Bayesian analysis of Kriging. *Technometrics*, 35(4):403–410. 17
- Heinonen, J. (2003).** Geometric embeddings of metric spaces. Volume 90 of Report. 104, 105
- Higdon, D. (1998).** A process-convolution approach to modelling temperatures in the North Atlantic Ocean. *Environmental and Ecological Statistics*, 5(2): 173–190. 53
- Journel, A. G. y Huijbregts, C. J. (1978).** *Mining Geostatistics*. Academic Press, London. 7, 13, 17

- Kennedy, M. C. y O’Hagan, A. (2001).** Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):425–464. 93
- Kitanidis, P. K. (1983).** Statistical estimation of polynomial generalized covariance functions and hydrologic applications. *Water Resources Research*, 19(4):909–921. 17
- Krige, D. G. (1951).** A statistical approach to some basic mine valuation problems on the witwatersrand. *Journal of the Chemical, Metallurgical and Mining Society of South Africa*, 52(6):119–139. 7
- Krivoruchko, K. y Gribov, A. (2002).** Geostatistical interpolation and simulation with non-Euclidean distances. En Sanchez-Villa, X., Carrera, J., y Gomez-Hernandez, J. J., editores, *geoENV IV*, páginas 331–342. Kluwer Academic Publishers. URL <http://training.esri.com/campus/library/Bibliography/RecordDetail.cfm?ID=30588&startrow=1&hidpage=1&browseonly=0>. 53, 207
- Le, N. D. y Zidek, J. V. (2006).** *Statistical Analysis of Environmental Space-Time Processes*. Springer Series in Statistics. Springer, New York, 1.^a edición. 15
- Lindgren, F., Rue, H., y Lindström, J. (2011).** An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial

- differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(4):423–498. [4](#), [119](#), [187](#), [189](#), [204](#), [206](#), [208](#)
- Little, L. S., Edwards, D., y Porter, D. E. (1997).** Kriging in estuaries: As the crow flies, or as the fish swims? *Journal of Experimental Marine Biology and Ecology*, 213:1–11. [53](#), [63](#)
- Løland, A. y Høst, G. (2003).** Spatial covariance modelling in a complex coastal domain by multidimensional scaling. *Environmetrics*, 14(3):307–321. [53](#)
- Matérn, B. (1986).** *Spatial Variation*. Springer-Verlag Berlin and Heidelberg GmbH & Co. K, 2.^a edición. [17](#)
- Matheron, G. (1963).** Principles of geostatistics. *Economic Geology*, 58(8): 1246–1266. [7](#), [38](#)
- Moore, E. H. (1935).** *General analysis*, volume 1 of *Memoirs of the American Philosophical Society*. The American Philosophical Society, Philadelphia. [105](#)
- Nash, J. (1956).** The imbedding problem for Riemannian manifolds. *The Annals of Mathematics*, 63(1):20–63. [101](#)
- Neteler, M. y Mitasova, H. (2002).** *Open Source GIS: A GRASS GIS Approach*. The Kluwer international series in Engineering and Computer Science (SECS): Volume 689. Kluwer Academic Publishers, Boston, Dordrecht, London. [70](#), [219](#)

- Okabe, A., Satoh, T., y Sugihara, K. (2009).** A kernel density estimation method for networks, its computational method and a GIS-based tool. *International Journal of Geographical Information Science*, 23(1):7–32. 53
- Open Source Geospatial Foundation (en progreso).** Sistemas de información geográfica. URL http://wiki.osgeo.org/wiki/Libro_SIG. 61
- Parlamento y Consejo Europeos (2002).** Directiva 2002/49/CE del parlamento europeo y del consejo, de 25 de junio de 2002, sobre evaluación y gestión del ruido ambiental. URL <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32002L0049:ES:NOT>. 43, 219
- Pérez Ortiz, J. A. (1999).** Diccionario urgente de estilo científico del español. URL <http://www.dlsi.ua.es/~japerez/pub/pdf/duce1999.pdf>. 5
- Pourahmadi, M. (1999).** Joint Mean-Covariance models with applications to longitudinal data: Unconstrained parameterisation. *Biometrika*, 86(3). 181
- Pekalska, E. y Duin, R. P. W. (2005).** *The Dissimilarity Representation for Pattern Recognition: Foundations And Applications*. Machine Perception and Artificial Intelligence. World Scientific Publishing Company. 137
- R Development Core Team (2008).** *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org>. 70

- R Development Core Team (2009)**. *R Instalation and Administration*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://cran.r-project.org/doc/manuals/R-admin.pdf>. 238
- Rao, C. R. (1972)**. Estimation of variance and covariance components in linear models. *Journal of the American Statistical Association*, 67(337):112–115. 17
- Rao, C. R. (1979)**. MINQE theory and its relation to ML and MML estimation of variance components. *Sankhyā: The Indian Journal of Statistics, Series B*, 41(3/4):138–153. 17
- Rathbun, S. L. (1998)**. Spatial modelling in irregularly shaped regions: Kriging estuaries. *Environmetrics*, 9:109–129. 53
- Ribeiro, P. J. y Diggle, P. J. (2001)**. geoR: a package for geostatistical analysis. *R-NEWS*, 1(2):14–18. 70
- Rudin, W. (1962)**. *Fourier Analysis on Groups*. Inter-Science. 116
- Rudin, W. (1991)**. *Functional Analysis*. McGraw-Hill, 2.^a edición. 115, 119, 205
- Rue, H. y Held, L. (2005)**. *Gaussian Markov Random Fields: Theory and Applications (Chapman & Hall/CRC Monographs on Statistics & Applied Probability)*. Chapman and Hall/CRC, 1.^a edición. 187, 189
- Rue, H. y Tjelmeland, H. (2002)**. Fitting Gaussian Markov random fields to Gaussian fields. *Scandinavian Journal of Statistics*, páginas 31–49. 189

- Rue, H., Martino, S., y Chopin, N. (2009).** Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71 (2):319–392. [187](#)
- SAGA Development Team (2008).** *System for Automated Geoscientific Analyses (SAGA GIS)*. Germany. URL <http://www.saga-gis.org/>. [70](#)
- Sasvári, Z. (1994).** *Positive Definite and Definitizable Functions*. Akademie Verlag. [109](#), [115](#)
- Schlather, M. (1999).** Introduction to positive definite functions and to unconditional simulation of random fields. Technical Report ST-99-10, Department of Mathematics and Statistics, Lancaster University, UK. URL <http://www.stochastik.math.uni-goettingen.de/~schlather/#Papers>. [17](#)
- Schlather, M. y Gneiting, T. (2006).** Local approximation of variograms by covariance functions. *Statistics & Probability Letters*, 76(12):1303–1304. [12](#)
- Schoenberg, I. J. (1937).** On certain metric spaces arising from Euclidean spaces by a change of metric and their imbedding in Hilbert space. *The Annals of Mathematics*, 38(4):787–793. [130](#)
- Schoenberg, I. J. (1938a).** Metric spaces and positive definite functions. *Transactions of the American Mathematical Society*, 44(3):522–536. [108](#), [110](#)
- Schoenberg, I. J. (1938b).** Metric spaces and completely monotone functions. *The Annals of Mathematics*, 39(4):811–841. [25](#), [26](#), [105](#), [110](#), [115](#)

- Stein, M. L. (1999).** *Interpolation of Spatial Data: Some Theory for Kriging* (Springer Series in Statistics). Springer, 1.^a edición. [17](#), [21](#), [26](#)
- Strichartz, R. S. (1974).** Fourier transforms and Non-Compact rotation groups. *Indiana University Mathematics Journal*, 24(6):499–526. [119](#), [177](#), [206](#)
- Tobler, W. R. (1970).** A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46:234–240. [8](#)
- Torgerson, W. S. (1952).** Multidimensional scaling: I. theory and method. *Psychometrika*, 17(4):401–419. [122](#)
- Torgerson, W. S. (1958).** *Theory and Methods of Scaling*. Wiley. [122](#)
- Ver-Hoef, J. M., Cressie, N., y Barry, R. P. (2004).** Flexible spatial models for kriging and cokriging using moving averages and the fast fourier transform (FFT). *Journal of Computational & Graphical Statistics*, 13(2): 265–282. [53](#)
- Ver-Hoef, J. M., Peterson, E. E., y Theobald, D. (2006).** Spatial statistical models that use flow and stream distance. *Environmental and Ecological Statistics*, 13(4):449–464. [208](#)
- Waller, L. A. y Gotway, C. A. (2004).** *Applied Spatial Statistics for Public Health Data*. Wiley-Interscience, 1.^a edición. [56](#)

- Watson, G. N. (1944).** *A Treatise on the Theory of Bessel Functions*. Cambridge Mathematical Library. Cambridge University Press, 2.^a edición. 34, 35
- Weisstein, E. W. (2010).** Poisson integral. From MathWorld—A Wolfram Web Resource. URL <http://mathworld.wolfram.com/PoissonIntegral.html>. 34
- Wolfram Research, Inc. (2008).** *Mathematica v.7.0*. Wolfram Research, Inc., Champaign, Illinois. 36, 157
- World Health Organization (1999).** Guidelines for community noise. Outcome of the WHO-expert task force meeting. URL <http://www.who.int/docstore/peh/noise/guidelines2.html>. 42
- Wurm, A., Krausz, N., DeWitt-Morette, C., y Berg, M. (2003).** Fourier transforms of Lorentz invariant functions. *Journal of Mathematical Physics*, 44(1):352–365. 166
- Yaglom, A. M. (1987).** *Correlation Theory of Stationary and Related Random Functions: Volume I: Basic Results*. Springer, 1.^a edición. 20

