

# NOTICIAS Y COMENTARIOS

---

## LA MATRIZ GEOGRÁFICA, CASI TREINTA AÑOS MÁS TARDE<sup>1</sup>

Muy probablemente fuera Berry<sup>2</sup> el primero en emplear el término *matriz geográfica*, para referirse a una tabla de doble entrada que representaba los valores de una serie de características (*filas*) en unidades espaciales determinadas (*columnas*).

En versiones posteriores, no obstante, lo habitual es que cada fila de la matriz geográfica represente una *unidad espacial* y cada columna, una muestra de una *característica* específica. Esta divergencia no tiene mayor importancia: en la representación habitual se prima la inspección visual de la matriz, que normalmente se hará unidad tras unidad espacial; en la representación de Berry, en cambio, se organizan los datos de cara a su procesado geoestadístico más frecuente (que es el que considera que los objetos de estudio son las unidades espaciales, no las características).

Desde el primer momento, los términos *casos* y *variables* (habituales en estadística) fueron incorporados al lenguaje geográfico cuantitativo como sinónimos de unidades espaciales y de características, respectivamente.

A la recogida de información seguía un análisis estadístico que pretendía, por un lado, reducir el número de variables y, por otro, reunir los casos en grupos homogéneos, o en conjuntos funcionales. El objetivo era simplificar la matriz original, recortando el número de variables (mediante

---

<sup>1</sup> La versión definitiva de esta nota recoge muchos comentarios y sugerencias de Joaquín Bosque, Emilio Chuvieco, Severino Escolano, Michael Gould, Pedro Reques y Vicente Rodríguez.

<sup>2</sup> BERRY, B. J. L. (1964): «Approaches to Regional Analysis: a Synthesis», *Annals of the Association of American Geographers*, pp. 2-11.

supresión de covariables, o elaboración de supervariables: componentes principales o factores) y el número de casos (mediante clasificación).

La matriz geográfica adquiere volumen, si se considera también la dimensión temporal. Idealmente la matriz se convierte en un paralelepípedo; en la práctica, sin embargo, dado que no se puede registrar el paso del tiempo de manera continua y sí de manera discreta, el resultado es una sucesión de matrices fechadas.

Cada matriz fechada es como una instantánea de la realidad. El intervalo entre instantánea es, normalmente, el mismo. Ello lleva consigo que la dimensión temporal de algunos sucesos quede distorsionada o, incluso, anulada. Habitualmente, los momentos inicial o final de los sucesos no coinciden con la fecha de una de las instantáneas. Además, hay sucesos que pueden ocurrir en el intervalo que media entre dos instantáneas. Es el problema de la frecuencia de todo muestreo, de su resolución, de la relación entre ésta y la variabilidad del fenómeno en cuestión.

En una sucesión de matrices fechadas, suponiendo que los casos y las variables no se alteren, se distinguen tres dimensiones ortogonales: la espacial (casos), la temática (variables) y la temporal. En ese espacio, el movimiento puede adoptar muy distintos patrones: el movimiento a lo largo de una arista del paralelepípedo es unidimensional (espacial —«vertical»—, temático —«horizontal»— o temporal —«en profundidad»—; el movimiento sobre una cara del sólido es bidimensional (temáticoespacial, temáticotemporal o espaciotemporal); el movimiento en el interior del poliedro es tridimensional (temáticoespaciotemporal).

A principios de los 60 se produjo el primer contacto entre la Geografía y la Informática, a raíz de la creación del *Sistema de Información Geográfica del Canadá*.<sup>3</sup> Desde entonces, los términos de corte estadístico han sido sustituidos paulatinamente por sus equivalentes en el lenguaje de los informáticos.

Esta nueva terminología enlaza directamente con la de la teoría de conjuntos: elementos, clases, propiedades y relaciones. Los casos (unidades espaciales) de la matriz geográfica se convierten en los *elementos* del

---

<sup>3</sup> HERES, L. y DE WINTER, N. (1990): «Towards a Standardized European Road Network Database», *Transportation*, 17, pp. 301-312.

conjunto problema. En lugar de variables, se habla de *propiedades* o, cuando corresponde, de *relaciones*, que se utilizan en la *clasificación* de los elementos implicados.

En un momento posterior, la terminología informática cambia ligeramente, imponiéndose los términos *entidad*, *relación* y *atributo*,<sup>4</sup> cuya correspondencia con los conceptos previos no requiere explicación. Calkins y Marble<sup>4</sup> publican en 1987 el primer artículo de gran difusión sobre la aplicación del modelo entidad-relación al diseño de un sistema de información geográfica orientado a la producción de cartografía vial.

Más recientemente todavía, hemos asistido a otro cambio en la terminología informática, patrocinado por los partidarios de una metodología de programación y de organización de bases de datos para *objetos*,<sup>4</sup> que utilizan indistintamente los términos propiedad, característica e, incluso, atributo. También se contemplan las relaciones entre objetos, aunque el énfasis se ponga siempre en éstos, no en aquéllas. No obstante, hay dos tipos de relaciones que son tenidas muy en cuenta: las relaciones de *generalización* y las de *agregación*. Existe relación de generalización entre dos objetos cuando uno de ellos, por ejemplo ferrocarril, es una especie del otro, medio de transporte, en este caso.<sup>5</sup> Existe relación de agregación

---

HERES, L. y WODD, T. F. (1992): «GDF, A Lingua Franca for Geographic Information», *EURONAV92 Proceedings*, November 17-19, London, 7 pp.

MASON, D.; O'CONNILL, M. y BELL, S. (1992): «Towards a 4-D GIS for Environmental Modelling», *Mapping Awareness & GIS in Europe*, vol. 6, n. 6, pp. 42-45.

TOMLINSON, R. F. (1984): «Geographic Information Systems —A New Frontier», Keynote Address to the *International Symposium on Spatial Data Handling*, August 20-24, Zurich, Switzerland.

WOOD, T. F. (1993): «How Europe is Addressing the Standards Issue», *GIS WORLD*, vol. 6, n. 1, pp. 42-47.

<sup>4</sup> CALKINS, H. W. y MARBLE, D. F. (1987): «The Transition to Automate Production Cartography: Design of the Master Cartographic Database», *The American Cartographer*, vol. 14, n. 2, pp. 105-119.

CHEN, E. F. (1970): «A Relational Model of Data for Large Shared Data Banks», *Communications of the ACM*, vol. 13, n. 6, pp. 377-387.

FRANK, A. U. y EGENHOFER, M. J. (1988): *Objet-Oriented Database Technology for GIS*, University of Maine, Dept. of Surveying Engineering, Orono, ME 04469, 76 pp.

<sup>5</sup> El concepto de generalización en Geografía tiene una tradición secular. La novedad a que nos referimos aquí es a la de su uso habitual en los medios informáticos, y consecuentemente en los medios SIGeográficos, para referirse a ese tipo de procedimiento de abstracción de datos que consiste en la búsqueda de puntos de convergencia entre objetos distintos, a pesar de las peculiaridades que los distinguen.

entre las parte y el todo (verbigracia, entre tramos de carretera y carreteras). Por otra parte, no se considera completa la definición de un objeto, si no se especifican las propiedades dinámicas del mismo, su comportamiento, es decir, qué procedimientos o *métodos* pueden en buena lid manipularlo.

Llegamos así a la reciente formulación del estándar de intercambio de información geográfica (espacial, en el original), encomendada a diversos comités y grupos de trabajo coordinados por el U.S. Geological Survey.<sup>6</sup> El estándar (SDTS) es un principio de acuerdo entre todos los colectivos que usan habitualmente información geográfica, entre ellos, muchos organismos gubernamentales. Para lograr dicho acuerdo no hay más remedio que definir inequívocamente los fundamentos de dicha información. En las definiciones del estándar se reflejan la mayoría de los términos cuya evolución estamos rastreando.

Parte el estándar de la consideración de tres conjuntos espaciales: *el conjunto de los fenómenos, el de las entidades y el de los objetos*.

Se considera fenómeno espacial todo hecho u ocurrencia externa. Los fenómenos son clasificados y caracterizados por el intelecto humano, dando lugar a distintas clases de entidades, definidas por una lista de atributos única y, en algunos casos, por una relación de criterios discriminantes de las otras clases de entidades. Por lo tanto, una clase de entidades, o entidad tipo, es el resultado del conocimiento de un grupo de fenómenos semejantes; una entidad (espacial) determinada es siempre un miembro de una clase (conjunto) de entidades.

A continuación se plantea la representación digital de las entidades espaciales, que se lleva a cabo mediante uno o varios objetos espaciales. El estándar define objetos espaciales elementales y objetos espaciales compuestos; objetos estrictamente geométricos (con localización y forma) y objetos topológicos (con localización, forma y vecindad). Entre los

---

<sup>6</sup> FEGEAS, R. G.; CASCIO, J. L. y LAZAR, R. A. (1992): «An Overview of FIPS 173, The Spatial Data Transfer Standard», *Cartography and GIS*, vol. 19, n. 5, pp. 278-293.

GREENLEE, D. D. (1992): «Developing a Raster Profile for the Spatial Data Transfer Standard», *Cartography and GIS*, vol. 19, n. 5, pp. 300-302.

LAZAR, R. A. (1992): «The SDTS Topological Vector Profile», *Cartography and GIS*, vol. 19, n. 5, pp. 296-299.

U. S. Geological Survey (1991): *Spatial Data Transfer Standard*, versión 12/90, Reston, Virginia, 320 pp.

geométricos se encuentran el punto, el segmento lineal, la cuerda, el arco, el polígono geométrico, el pixel y la malla unidad. Entre los topológicos, el nodo, la cadena, el polígono topológico y el polígono universal.

En la terminología del estándar se denomina *feature* al par (entidad espacial, objeto/s), para evitar la confusión que aparentemente existía cuando no se diferenciaba una entidad de su representación, pues lo normal es que se las denominara *features* a ambas. Esta distinción es especialmente importante cuando, como es habitual, una entidad es representada por más de un objeto. En este caso, la persona que considera una entidad se desentiende, como debe ser, de su representación interna.<sup>7</sup>

El estándar americano está dividido en tres secciones fundamentales. En la primera se aborda la presentación de los conceptos y la estructura general de la información a transmitir. En la segunda parte se aborda la definición de 200 *features* tipo, relacionadas con las descripciones topográficas e hidrológicas, la definición de 244 atributos y el listado de 1.250 términos más, que se consideran sinónimos o especificaciones de los 444 términos definidos previamente. En la tercera parte se explica la codificación física de la información a transmitir, en cualquiera de los soportes contemplados por las normas ISO/ANSI 8211.

Para definir la estructura de los datos a transmitir el estándar ofrece 34 tipos de módulos distintos. En la mayoría de las transferencias no es necesario recurrir más que a un subconjunto de ellos. Los distintos módulos permiten registrar información global acerca de los datos y de su calidad, el diccionario de *features* y atributos, el sistema de coordenadas, y la localización, atributos y relaciones de los objetos.

Para facilitar el uso del estándar para intercambio de datos espaciales y sus relaciones topológicas se ha desarrollado un perfil específico (*Topological Vector Profile, TVP*), que es un subconjunto de las normas de transferencia globales.<sup>6</sup> El perfil de intercambio de información raster está todavía en período de prueba<sup>6</sup>; cuando se distribuya permitirá el intercambio flexible de modelos digitales del terreno, de ortofotos digitalizados y de imágenes digitales georeferenciadas.

Antes o después, por su carácter de estándar, se impondrá esta termi-

---

<sup>7</sup> LANGRAN, G. (1992): *Time in GIS*, London, Taylor & Francis, 189 pp.

nología. Por lo tanto, creemos, estos párrafos pueden ser muy útiles para comprender todas las formulaciones que recurran a ella. En pocas palabras, se puede decir que el estándar americano considera y *distingue* la realidad espacial (es decir, los fenómenos) de su conocimiento y estructuración por el intelecto humano en conjuntos de entidades. Este estándar, también, diferencia las entidades de su representación en un marco digital como combinaciones de objetos espaciales.

Aunque no se haya formulado todavía el estándar europeo equivalente, existen ya varios estándares de intercambio de datos espaciales de alcance regional y/o sectorial. Entre ellos cabe distinguir BS7567 (NTF), desarrollado en el Reino Unido, DIGEST, redactado por los servicios cartograficogeográficos de las fuerzas armadas de la OTAN, EDIGéo y GDF.<sup>3</sup> Este último ha sido desarrollado por un consorcio multinacional: Daimler Benz, Intergraph, Philips, Renault, Bosch, SAGEM, Teleatlas, MVA Systematica y el Instituto de Cartografía de la Universidad de Hannover. La primera versión de GDF fue publicada en 1988 y la segunda, en enero de 1992. Ya está anunciada una nueva versión de GDF para enero de 1995. GDF fue concebido originalmente como estándar de mapas digitales de carreteras, para sistemas de «navegación terrestre» a bordo de vehículos automóviles. El modelo conceptual de GDF procede de la misma fuente que el del estándar americano (SDTS), a saber, el modelo entidad-relación.<sup>4</sup> Por ello, sus diferencias no pueden ser importantes; por ejemplo: en GDF existen módulos específicos para representar las relaciones, en SDTS, en cambio, se utilizan los mismos módulos que para representar atributos.

GDF, como SDTS, contempla distintos niveles de descripción espacial, desde el meramente pictórico hasta el rigurosamente estructurado (topológico). En versiones futuras se piensa incluir en GDF una sección dedicada al tratamiento de la dimensión temporal. GDF, como SDTS (*vid. supra*), asume la norma ISO/ANSI 8211 de transmisión genérica de datos.

Terminamos estas páginas con una aclaración sobre la interpretación geográfica actual: *las matrices geográficas*, que son el objeto de esta nota, y *las variables espaciales bidimensionales*.<sup>8</sup> A pesar de su significación, no

---

<sup>8</sup> Otro tipo frecuente de matrices son las que reflejan las relaciones entre entidades espaciales: matrices cuadradas (no siempre simétricas) de distancias, flujos migratorios, intercambios económicos, matrices de contigüidad, etc.

nos consta por ahora que se haya escrito con rigor sobre este tema. Quien haya tenido que explicar estos conceptos, habrá constatado lo confuso que resulta para la mayoría de los geógrafos que se use el mismo término (matriz) al hablar de una tabla de doble entrada que al hacerlo de una imagen digital en sentido amplio (retículo ortogonal de valores numéricos correspondientes a la intensidad de un determinado fenómeno espacial).

Vayamos por partes. Se puede hablar de matrices en ambos casos porque el término matriz, a secas, significa únicamente formación en filas y columnas de elementos de la misma clase: nombres, rangos, números, etc.

Pero la matriz geográfica —versión habitual— es propiamente un conjunto de vectores (vectores fila, si nos interesan las unidades espaciales; vectores columna, si prestamos atención a las características). Si la matriz en cuestión tiene  $n$  variables y  $m$  casos, la información que reúne es  $n$  dimensional, tantas dimensiones como variables (o  $m$  dimensional, tantas como casos). Si se quiere tratar la información que reúnen dos o más columnas de la tabla, habrá que recurrir a algún procedimiento multidimensional. Una matriz geográfica es el resultado de muestrear  $n$  variables en  $m$ , siempre los mismos, lugares. Cada columna de esa matriz es, pues, un plano temático. Consecuentemente, el conjunto de todas las columnas de la matriz es la superposición (*overlay*) de  $n$  capas temáticas.

Una imagen digital, en cambio, es la muestra de una única variable en puntos discretos del espacio bidimensional con una distribución regular, ortogonal. Filas y columnas no tienen otro significado que el de perfiles este-oeste, o norte-sur, en la superficie, única, de variación. Se requieren dos imágenes, por lo menos, para efectuar un análisis multidimensional (cfr. sistemas de tratamiento digital de imágenes). Cada imagen es solamente un plano temático.

Se comprende ahora que los tratamientos a que se someten los dos tipos distintos de matrices sean diferentes. La matriz geográfica, como el paisaje complejo que representa, es claramente multidimensional. No lo es en cambio la imagen, digital, de alguno de sus componentes aislados.

La dimensionalidad de los datos geográficos varía según el objeto de estudio. En el caso de considerar solamente la situación planimétrica de los individuos y una característica nos encontramos ante un problema tridimensional (más características añadirán nuevas dimensiones). Si consideramos también la altitud y el tiempo, por qué no, nos encontramos

ante un problema de cuatro dimensiones (de cinco, al considerar cualquier característica<sup>3</sup>). Si, en vez de interesarnos por un elemento geográfico puntual, trabajamos con elementos extensos, la dimensionalidad del problema se dispara.<sup>7</sup>

Juan A. CEBRIÁN

### **LAS FUENTES ORALES: SU UTILIDAD EN ESTUDIOS SOBRE MIGRACIONES**

Los estudios de geografía de la población con cierta frecuencia se han reducido en nuestro país al análisis de aquellos factores que contasen con aparato estadístico oficial, aunque éste en demasiadas ocasiones no fue ni muy abundante ni excesivamente fiable. A veces se constata como en numerosos estudios el apartado sobre migraciones se reducía a poco más que cifras (en lo referente a los saldos migratorios), a un recuento de los lugares de origen y alguna otra característica de los inmigrados.

Afortunadamente, conforme nos acercamos al momento actual, los estudios geográficos sobre migraciones han ido ganando en variedad de factores analizados y en profundidad en el tratamiento de los mismos, al amparo de la indudable mejora de los datos estadísticos disponibles, del mayor bagaje metodológico de los geógrafos actuales, de las aportaciones informáticas y del desarrollo de técnicas paralelas de acceso a datos de interés no recogidos oficialmente (como las encuestas).

Otro de los medios que en los últimos años comienzan a ser utilizados por los estudiosos de los movimientos migratorios, aunque de modo todavía muy minoritario, son las denominadas fuentes orales, tratadas actualmente con un rigor científico comparable al de la mayoría de las otras técnicas paralelas de acceso a datos de interés no recogidos oficialmente (como las encuestas).

Otro de los medios que en los últimos años comienzan a ser utilizados por los estudiosos de los movimientos migratorios, aunque de modo todavía muy minoritario, son las denominadas fuentes orales, tratadas actual-