## *LA CORRELACION LINEAL:*

### *Precisiones prácticas y su funcionalidad en la determinación de las similitudes y diferencias de los espacios geográficos*

JOSE L. GURRIA GASCON

La correlación bidimensional es una operación matemática que mide, a través de su índice, una "dependencia", "relación recíproca" o "asociación" entre dos variables [1].

En ninguno de los estudios consultados hemos visto que se hiciera referencia o se incluyera como técnica en la definición de espacios geográficos. De acuerdo con Buckley [2], los espacios geográficos se definen tanto por sus similitudes como por sus diferencias con otros, fundamentalmente con los más próximos.

Desde este punto de vista, estamos en la idea —y así pretendemos demostrarlo en este artículo— de que la correlación simple —con un pequeño manejo— es una técnica matemática válida para la determinación de similitudes y diferencias entre sistemas o subsistemas, en definitiva, espacios geográficos.

1. Díaz Alvarez, Jr. (1977), «Notaciones sobre la metodología del análisis cuantitativo aplicado a la Geografía», *Paralelo 37º,* 1, Almería; Compán Vázquez, D. (1978), «Sobre el uso de la correlación lineal simple en Geografía. Aplicación al estudio de la distribución espacial de la renta en España», *Cuadernos Geográficos,* 8, Granada; Estébanez Alvarez, J. y Bradshaw, P. S. (1979), *Técnicas de cuantificación en Geografía,* Tebar Flores, Madrid; Harvey, D. (1979), *Explanation in Geography,* E. Arnold, London; Grupo Chalude (1980), *Iniciación a los métodos estadísticos en Geografía,* Ariel, Barcelona; Hammond, R. y McCullagh, P. S. (1980), *Técnicas cuantitativas en Geografía,* Saltés, Madrid; Unwin, D. (1981), *Introductory spatial analysis,* Methuen, London.
2. Buckley, W. (1973), *La sociología y la teoría moderna de los sistemas,* Amorrortu, Buenos Aires (2.ª edi.). En la p. 16, este autor afirma textualmente que «la meta principal del movimiento de investigación de sistemas generales es, pues, perfilar esas semejanzas estructurales, y al mismo tiempo, distinguir las diferencias estructurales entre sistemas de tipos sustancialmente distintos».

Esta nueva proyección de la correlación se desarrolló, en principio, en nuestra Tesis Doctoral [3], con la intención de aplicarla en la delimitación de la montaña extremeña. Sin embargo, en aquella ocasión, quedó un poco al margen, ya que, desbordados por la vasta matriz de datos, fue necesario aplicar, para tal fin, un análisis factorial.

Es esta una de las principales limitaciones de la correlación lineal: no puede competir en la definición y delimitación de espacios geográficos con los análisis multivariantes, ni puede hacer frente a una vasta matriz de datos. Sólo es posible analizar dos variables conjunta y simultáneamente.

Por otra parte, han de extremarse, como se verá más adelante, todas las precauciones tanto en la expresión estadística de las variables como en la interpretación final de sus resutlados, aunque estas notas deben hacerse extensibles a cualquier operación matemático-estadística.

Por el contrario, aun teniendo en cuenta estos inconvenientes, las posibilidades que presenta la convierten en una operación de la que se pueden obtener buenos resultados en Geografía.

En primer lugar, es una técnica muy elemental y sencilla, que se puede realizar con una simple calculadora manual.

La fórmula que hemos aplicado a todos los ejemplos del trabajo, una vez desarrollada, es la siguiente:

$$R = \frac{N\Sigma XY - \Sigma X \Sigma Y}{\sqrt{\left[N\Sigma X^2 - (\Sigma X)^2\right]\left[N\Sigma Y^2 - (\Sigma X)^2\right]}}$$

En segundo lugar, puede explicar un alto porcentaje de la varianza de algunos fenómenos geográficos elementales, con una estructura sencilla, en la que el mayor peso lo ostentan dos o tres variables. Es el caso, por ejemplo, del aumento de las precipitaciones en áreas de montaña. La correlación entre las precipitaciones y el aumento de la altitud dará un alto índice, a la vez que explicará un elevado porcentaje de la varianza de este fenómeno.

Para constatar el porcentaje de la varianza explicado, basta con aplicar al índice de la correlación el Coeficiente de Determinación $= r^2 . 100$.

Remitimos, para un mayor detalle acerca de la correlación, a cualquiera de las obras ya reseñadas, y en especial la de Compán Vázquez

3. Gurria Gascón, J. L. (1983), *El paisaje de montaña en Extremadura* (delimitación, economía y población), setiembre 1983, Facultad de Filosofía y Letras de la Universidad de Extremadura, Cáceres (en imprenta).

o la de Estébanez Alvarez y Bradshaw [4], que nosotros seguiremos fundamentalmente para las cuestiones de tipo general y terminológico.

Finalmente, su índice puede aprovecharse asimismo, para la determinación de similitudes y diferencias entre espacios geográficos. Es el objetivo fundamental de este artículo, del que pasaremos a ocuparnos a continuación.

Los índices —como se sabe— pueden oscilar entre 0 y $\pm$ 1,000. Las distintas opiniones consultadas no están de acuerdo en los límites a partir de los cuales estos índices comienzan a ser significativos. Pero, a tenor de ello, se puede considerar en general que comienzan a ser significativos a partir del $\pm$ 0,300 [5] y son ya bastante considerables a partir del $\mp$ 0,497 [6].

Cuando estos índices son bajos, se habla de "independencia aleatoria". Quiere decir que las dos variables entre sí guardan un comportamiento arbitrario, y que el fenómeno o la estructura geográfica que se trata de estudiar debe explicarse por otras relaciones.

Para que el índice sea significativo es necesario que los valores de las dos variables experimenten una progresión aritmética. Si una de las variables presenta —como en el ejemplo siguiente— unos valores homogéneos a lo largo de toda la matriz, el índice será de 0,000

| | X | Y | X | Y |
|---|---|---|---|---|
| 1º | 12 | 10 | 5 | 7 |
| 2º | 11 | 10 | 5 | 7 |
| 3º | 9 | 10 | 5 | 7 |
| 4º | 6 | 10 | 5 | 7 |
| 5º | 5 | 10 | 5 | 7 |
| 6º | 3 | 10 | 5 | 7 |

En estos dos casos, un índice de 0,000 no deja, por ello, de ser sumamente significativo, por cuanto que implica una total similitud de los seis elementos considerados, al menos en tres de las cuatro variables expresadas. Sus diferencias, si existen, habrían de buscarse en otras variables.

Pero pongamos un ejemplo real y más tangible, sin ser tan extremo. Por la literatura existente sobre áreas de montaña, sabemos que ésta tiene una vocación predominantemente silvo-pastoril. Al realizar la correlación con las variables altitud y tierras forestales [7], incluyendo todos aquellos municipios (96 en total en la provincia de Cáceres) con alineaciones motañosas, el índice fue de 0,078. Un índice

4. Estébanez Alvarez, J. y Bradshaw, R. P. (1979), op. cit.; Compán Vázquez, D. (1978), op. cit.

5. Fernández Gutiérrez, F. (1979), «Consideraciones metodológicas y experimentales del análisis factorial en Geografía», *Cuadernos Geográficos*, 8, Granada.

6. García Ramón, María D. (1981), *Métodos y conceptos en Geografía rural*, Oikos-Tau, Barcelona.

7. Por razones de espacio no se incluye toda la vasta matriz de datos. Para su consulta, remitimos a nuestra Tesis Doctoral, ya mencionada.

insignificante, como lo fue el de la altitud y tierras pastadas (0,241).

¿Quería esto decir que la montaña extremeña no tenía vocación silvo-pastoril y que no encajaba en el contexto general de la montaña mediterránea, en la que el aumento de la altitud implicaba también el aumento de las tierras de pastos y bosques? Recurriendo a la matriz de datos, enseguida se pudo comprobar todo lo contrario: tenía, de hecho, esta vocación silvo-pastoril. Sin embargo, estos bajos índices eran evidentes, por cuanto que existía una homogeneidad en los valores de las dos segundas variables a lo largo de toda la muestra. Estos bajos índices estaban implicando una gran similitud del espacio montañés en cuanto a su vocación silvo-pastoril.

A continuación se seleccionaron por muestreo una serie de municipios propiamente de la penillanura, para ver el efecto que producían en los índices de correlación de estas mismas variables. El total de municipios era entonces de 111. En la primera correlación el índice (0,175) siguió siendo inapreciable, al igual que en la segunda, en la que incluso fue de signo negativo (—0,225).

Era también evidente, por cuanto que la penillanura cacereña, aunque de otro espacio geográfico se trate, tiene esta misma vocación silvo-pastoril. Así, aunque por distintas razones, toda la muestra seguía siendo homogénea y, en consecuencia, ni la variable tierras forestales, ni la variable pastos pueden mostrar "dependencia" o "asociación" con respecto a ninguna otra, ya que sus valores ni aumentan ni disminuyen a medida que lo pueda hacer cualquier otra variable.

Estos bajos índices de correlación nos están mostrando, por lo tanto, similitudes entre dos espacios geográficos fundamentalmente distintos.

Además ,estos índices nos habían descubierto una importantísima hipótesis de trabajo: la montaña extremeña, antes que insertarla en el contexto general de la montaña mediterránea, había que encajarla en su contexto regional.

Sin embargo, unos índices de correlación bajos no siempre implican similitud. El juicio del investigador, el recurso a la matriz de datos, a la cartografía y literatura existentes, así como a la labor de campo, deberá prevalecer. Ya se ha visto anteriormente cómo también pueden deberse "a un comportamiento arbitrario" de las dos variables entre sí.

Y, desde luego, pueden ser el resultado de anomalías en la matriz de datos. Estas anomalías suelen provenir o bien de una errónea cuantificación de las variables, o bien de la complejidad y excepcionalidad que presentan las mismas.

En el primer caso, como indica García Ramón [8], "cuanto más neutras sean las variables empleadas, menor es el riesgo de un posible

8. García Ramón, María D. (1981), op. cit.

error en la interpretación". Generalmente, la forma más objetiva es la cuantificación mediante la transcripción a datos porcentuales. Sin embargo, también es posible introducir otras medidas combinadas [9], como número de tractores/ha. de cultivo, UGM/ha. o por cada 100 has., etcétera.

En el ejemplo que se expone a continuación puede observarse cómo no es necesario la cuantificación por medio de porcentajes para que pueda haber una "dependencia funcional", una correlación perfecta ($\mp$ 1,000):

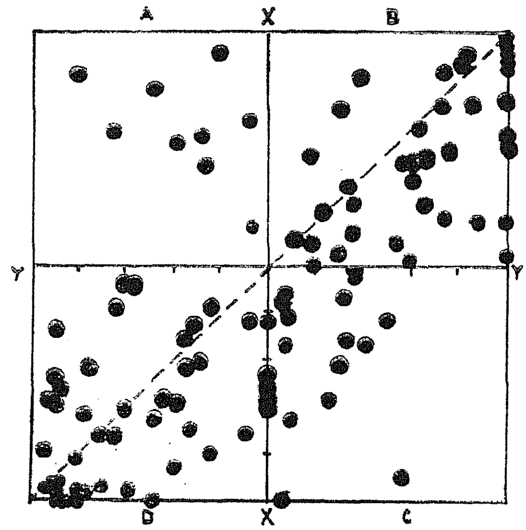
| | X | Y | |
|---|---|---|---|
| 1.º | 20 % | 0,1 | |
| 2.º | 40 | 0,2 | |
| 3.º | 60 | 0,3 | R = 1,000 |
| 4.º | 80 | 0,4 | |

En el segundo caso, existen algunas variables que por su complejidad o casuística pueden dar lugar a anomalías.

Una correlación muy tangible se da normalmente entre la altitud y la pendiente: a mayor altitud-mayor pendiente. Sin embargo, al aplicarla a todos los municipios de la provincia de Cáceres, su ndice (0,411) no resultaba suficientemente expresivo. Pero recurriendo a los mapas geológicos, a los topográficos y a la foto aérea, enseguida se pudieron detectar una serie de anomalías, derivadas de la peculiar configuración morfogenética de la provincia.

9. Brocard, M., Pumain, D. y Rey, V. (1977), «L'nalyse de donnée traitements visuels et mathematiques», *L'Espace Geographique,* VI, 4, Paris. Estos autores, aunque refiriéndose al análisis en componentes principales, afirman que se pueden tratar datos ordinales (a partir de una matriz de correlación de rangos) e incluso matrices de ausencia-presencia.

Con los datos de estas dos variables (% de altitud por encima de los 500 m. y % de pendiente de valor superior al 20%, a nivel municipal) trazamos el "diagrama de dispersión" siguiente:




(X) ALTITUD / PENDIENTE (Y)

(Unicamente se han incluido algunos de los más significativos, aproximadamente el 50%).

Enseguida se pudieron "diferenciar" cuatro conjuntos perfectamente definidos:

— En el subcuadrante A se localizaban todos aquellos municipios con un alto porcentaje de altitud, pero muy bajo en cuanto a la pendiente. Son todos aquellos municipios enclavados en los bloques más elevados de la penillanura, bien por la acción de una tectónica más o menos reciente, o bien por diferencias morfogenéticas, producto de una erosión diferencial entre granitos y pizarras. Otros, fuera de la penillanura, se enmarcaban en los piedemontes del Sistema Central y Villuercas, con extensos glacis de erosión (Sistema Central) o de acumulación (caso de las rañas de Villuercas).

— Por el contrario, en el subcuadrante C se localizan aquellos municipios con fuertes pendientes, pero con escaso porcentaje de altitud por encima de los 500 m. Son los municipios de penillanura encuadrados en la zona de "riveros" del Tajo, donde el profundo encajamiento de la red fluvial sobre un sustrato metamórfico precámbrico muy endurecido produce estas fuertes pendientes.

En estos dos grupos de municipios habría que incluir otros de los subcuadrantes B y D, con alguna de las variables muy extremas. Son todos aquellos que están alejados de la diagonal.

Pero únicamente eliminando los subcuadrantes A y C, con lo que las dos variables quedaban más purificadas, el índice de la correlación alcanzó ya el valor de 0,830 [10].

Si los índices bajos podían reflejar —aunque no necesariamente— unas similitudes entre espacios geográficos, los índices altos marcan —muchas veces— además de una "dependencia" o "asociación", una diferenciación en cuanto a las dos variables correlacionadas.

Es evidente que en el ejemplo siguiente, con un índice de -1,000, existe una clara diferencia entre cada uno de los individuos, pero especialmente entre el 1.º y el 5.º:

| | X | Y |
|---|---|---|
| 1.º | 20 | 100 |
| 2.º | 40 | 80 |
| 3.º | 60 | 60 |
| 4.º | 80 | 40 |
| 5.º | 100 | 20 |

Pero para que exista realmente una diferenciación de espacios geográficos será necesario que esta "dependencia" entre las dos variables sea de causa-efecto. Las relaciones que ligan a dos variables entre sí pueden ser, según Aracil [11], causales o correlativas (por ejemplo, estadísticas). Para su determinación también ha de prevalecer el juicio y la opinión del investigador [12]. Puede aplicarse igualmente la prueba "t" de Student para determinar en qué grado se debe al azar el índice extraído de cualquier correlación.

Esta diferenciación se da en dos vertientes: en primer lugar, por la diferencia de signo del índice; y en segundo lugar, por las diferencias de valores que deben existir a lo largo de las dos variables, es decir, que necesariamente debe aumentar o disminuir la una, mientras la otra también lo hace.

10. Se había elegido la curva de nivel de los 500 m. por tener la penillanura una altitud media entre los 400 y 450 m., por lo que los relieves por encima de los 500 m. ya destacaban sobre su entorno. Sin embargo, si en vez de utilizar esta cota, se considerara a partir de los 600-650 m., se eliminarían casi todas las anomalías del subcuadrante A.

11. Aracil. J. (1978), *Introducción a la dinámica de sistemas,* Alianza, Madrid.

12. Hammond y McCullagh (1980), op. cit.

En el primer caso, el signo puede proporcionarnos un elemento diferenciador de estructuras y espacios geográficos.

El signo más (+) señala correlaciones directas: a medida que aumenta o disminuye una variable, también simultáneamente la otra. Este signo nos está indicando que las dos variables son características del mismo espacio o estructura.

El signo (—) es debido a una correlación inversa: a medida que aumenta o disminuye una variable, la otra lo hace en sentido inverso. En este caso, el signo está indicando que cada una de las variables corresponde a una estructura o espacio geográfico distinto.

Así, a través del signo, se pueden definir estructuras y espacios geográficos "por lo que son", es decir, por sus variables propias, y "por lo que no son", es decir, por las variables propias de otros espacios. Define, por lo tanto, en relación con otros espacios.

En el segundo caso, en cuanto a las diferencias de valores en las variables, hemos de decir que la correlación se muestra especialmente sensible a los individuos con valores extremos, que pueden introducir "graves anomalías". En concreto, en el ejemplo que se expone a continuación, el índice es de 0,774:

| | X | Y |
|---|---|---|
| 1.º | 45 | 34 |
| 2.º | 38 | 41 |
| 3.º | 36 | 37 |
| 4.º | 29 | 31 |
| 5.º | 28 | 32 |
| 6.º | 26 | 29 |
| 7.º | 21 | 24 |

El índice es ya bastante alto, pero introduciendo nada más que dos valores extremos:

| | |
|---|---|
| 15 | 12 |
| 10 | 8 |

el índice general aumenta considerablemente, nada menos que hasta 0,906.

Y si los valores son todavía más extremos, se alcanza casi una correlación perfecta, con 0,986:

| | |
|---|---|
| 98 | 90 |
| 99 | 98 |

Si estos dos últimos valores son más extremos no es debido a que numéricamente, cuantitativamente, tengan un mayor peso en las operaciones por ser valores más elevados, sino a que están más alejados de la serie expuesta anteriormente. Hay que tener en cuenta que un valor extremo provoca la misma anomalía tanto si es bajo como si es alto. En el cuadro siguiente, el índice de 0,118, insignificante, pero que resulta lógico, como consecuencia de un grupo homogéneo, se elevo hasta 0,996 al introducir un valor muy bajo (entre rayado):

| | | | |
|---|---|---|---|
| 1.º | 98 | 90 | |
| 2.º | 99 | 98 | |
| 3.º | 98 | 95 | |
| 4.º | 97 | 99 | R = 0,118 |
| 5.º | 96 | 94 | |
| 6.º | 98 | 97 | |
| | 3 | 5 | R = 0,996 |

Quedan reflejadas, por lo tanto, a través de distintos índices, unas "anomalías", unas "diferencias de proximidad" de unos valores con relación a un conjunto "base". Si estos valores extremos, en vez de ser individuos aislados, fueran también conjuntos homogéneos, no cabe duda de que quedarían marcadas las diferencias entre espacios geográficos distintos. Veamos un ejemplo real.

Si el índice entre la altitud y las tierras forestales o de pastos no resultaba significativo, en cambio el de altitud y cultivos herbáceos es de -0,648 en la provincia de Cáceres. Podemos concluir que existe una "dependencia" de causa y efecto entre estas dos variables (fundamentalmente en lo concerniente a los cereales). Teniendo esto en cuenta, se seleccionó un conjunto de municipios de penillanura y otro de montaña:

| | | |
|---|---|---|
| Malpartida de Plasencia ... ... ... ... ... | 4,9 | 46,0 |
| Oliva de Plasencia ... ... ... ... ... ... ... | 5,0 | 40,1 |
| Plasencia ... ... ... ... ... ... ... ... ... ... ... | 9,8 | 71,0 |
| Trujillo ... ... ... ... ... ... ... ... ... ... ... | 12,3 | 66,3 |
| Pozuelo de Zarzón ... ... ... ... ... ... ... | 15,2 | 69,2 |
| Jaraíz de la Vera ... ... ... ... ... ... ... | 13,4 | ¿? |

Puesto que se trata de un grupo más o menos homogéneo en cuanto a estas dos variables (escaso porcentaje de altitud por encima de los 500 m. y alto porcentaje destinado a cereales), sería imposible determinar que Jaraíz, con mayor altitud que Malpartida, tiene un 84,6% de herbáceos, al contrario de lo que indica el índice general de correlación. En este grupo sólo se podría determinar, por el cálculo de probabilidades, que a un bajo porcentaje de tierras por encima de los 500 m. corresponde un alto porcentaje de cultivos herbáceos. Como en la realidad las progresiones aritméticas no suelen realizarse de una forma perfecta, como en algunos ejemplos puestos anteriormente, el índice de correlación de este grupo parece ser y demostrar lo contrario que el índice general aplicado para toda la muestra de municipios.

En el caso siguiente, sería igualmente imposible determinar, por el cálculo de probabilidades, si Marchagaz tiene más o menos herbáceos que el resto de los individuos de su grupo, pero sí se podría afirmar que Marchagaz tiene menor porcentaje que cualquier individuo del conjunto anterior de la penillanura.

| | | |
|---|---|---|
| Ladrillar ... ... ... ... ... ... | 87,5 | 21,9 |
| Nuñomoral ... ... ... ... ... ... | 91,8 | 15,5 |
| Tornavacas ... ... ... ... ... ... | 100,0 | 21,7 |
| Jerte ... ... ... ... ... ... ... ... | 100,0 | 5,9 |
| Marchagaz ... ... ... ... ... ... | 100,0 | ¿? |

Al correlacionar sólo estos dos grupos conjuntamente, la correlación saldría mucho más elevada que la general (—0,648), por un efecto de "dependencia general". Así el índice es de —0,849.

Esto nos sugiere que al ir correlacionando grupos sucesivos, más o menos homogéneos, de la muestra con relación a un grupo "base" (el de penillanura), los coeficientes probablemente fueran marcando sucesivamente distintas dependencias, en relación con la proximidad o alejamiento a ese grupo "base", lo que no es sino una diferenciación "en relación a".

El problema radica en establecer los distintos conjuntos, más o menos homogéneos, e ir relacionándolos con un conjunto "base" sucesivamente y por separado.

Ha de realizarse, en primer lugar, la operación para toda la muestra de individuos, y si el índice es significativo y de causa-efecto, la correlación implicará unas diferencias entre los individuos. Estos se pueden agrupar entonces en conjuntos simplemente atendiendo a una sola variable. Es suficiente si está fuertemente correlacionada con otras.

Basta con agrupar los distintos municipios o individuos dividiéndolos en conjuntos que agrupen valores de 20 en 20, o de 10 en 10 si se persigue una mayor precisión.

Nosotros, en concreto, realizamos una división en función de la altitud y de la pendiente, por ser estas dos variables inherentes al concepto de montaña:

— Grupo A: municipios con un porcentaje superior al 80% por encima de los 500 m. y de pendiente con desniveles de más del 20%, o bien:

$$\frac{\%\ \text{altitud} + \%\ \text{pendiente}}{2} = \text{al } 80\% \text{ o superior.}$$

— Grupo B: municipios con porcentajes entre el 60 y el 80% o con una media entre las dos variables igual o superior al 60%.

— Grupo C: municipios entre el 40 y el 60% o con media igual o superior al 40%.

— Grupo D: municipios entre el 20 y el 40% o con media superior al 20%.

— Grupo E: municipios entre el 10 y el 20% o con media superior al 10%.

— Grupo F: es el conjunto base, con el que se irán correlacionando por separado cada uno de los restantes grupos. Está formado por municipios de penillanura, con porcentajes de altitud y pendiente inferiores al 10% [13].

De estas divisiones, resultaron 23 municipios en el Grupo A, 21 en el B, 24 en el C, 13 en el D, 7 en el E y 23 en el F.

Los índices de correlación resultantes constantaron la hipótesis, marcando gradualmente las diferencias. Las expresadas a continuación son las que se realizaron:

GRUPOS F/A

| | |
|---|---|
| — Altitud/Pendiente ... ... ... ... ... ... ... | 1,000 |
| — Altitud/Tierras de labor ... ... ... ... ... ... | —0,654 |
| — Altitud/Parcelas de —1 ha. ... ... ... ... | 0,652 |
| — Altitud/% encinar y alcornocal ... ... ... | —0,880 |

13. Como los distintos grupos habían salido con un número total de municipios muy irregular, por temor a que los conjuntos con mayor número pudieran introducir anomalías, se iban eliminando individuos del grupo base hasta conseguir un número igual al del otro conjunto con el que se correlacionaba. No presentaba problemas ni con el Grupo A, ni con el B, ni con el C, que tenían un número igual o muy próximo. En principio, era suficiente para constatar la hipótesis.

Sin embargo, posteriormente hemos comprobado que un conjunto, aunque cuente con un número mayor o menor de individuos, no introduce apenas anomalías. En el ejemplo siguientes, se puede constatar cómo los índices de correlación apenas varían aunque uno de los conjuntos cuente incluso con el doble de individuos.

| | | |
|---|---|---|
| 45 | 34 | |
| 38 | 41 | |
| 36 | 37 | |
| 29 | 31 | |
| 28 | 32 | |
| 26 | 29 | |
| 21 | 24 | |
| | | = 0,774 |
| 98 | 90 | |
| 99 | 98 | |
| | | = 0,986 |
| 98 | 95 | |
| 97 | 99 | |
| 96 | 94 | |
| 98 | 97 | |
| | | = 0,992 |
| 93 | 97 | |
| 96 | 96 | |
| 95 | 96 | |
| 94 | 97 | |
| 95 | 95 | |
| 98 | 94 | |
| | | = 0,992 |

GRUPOS F/B

| | |
|---|---|
| 1.ª ... | 0,869 |
| 2.ª ... | —0,459 |
| 3.ª ... | 0,487 |
| 4.ª ... | —0,762 |

GRUPOS F/C

| | |
|---|---|
| 1.ª ... | 0,703 |
| 2.ª ... | —0,320 |
| 3.ª ... | 0,520 |
| 4.ª ... | —0,641 |

GRUPOS F/D

| | |
|---|---|
| 1.ª ... | 0,702 |
| 2.ª ... | —0,060 |
| 3.ª ... | 0,261 |
| 4.ª ... | —0,315 |

GRUPOS F/E

| | |
|---|---|
| 1.ª ... | 0,709 |
| 2.ª ... | —0,066 |
| 3.ª ... | —0,308 |
| 4.ª ... | —0,322 |

Pero quizás inseguros todavía de estos resultados, se aplicó la misma operación a aquellos conjuntos que habían salido del análisis factorial.

Indudablemente, el "grupo base" de penillanura había resultado muy definido como tal, y así sirvió de nuevo para tal fin. Pretendíamos constatar si la correlación también marcaba de alguna manera las diferencias establecidas por el análisis factorial en sus variables más definitorias. En este caso, ya no tuvimos en cuenta los conjuntos más próximos a la penillanura y nos basamos fundamentalmente en los conjuntos más montañosos, en un primer momento.

De esta manera, contábamos con un primer grupo (1), el más montañoso; un segundo grupo (2), de montaña media-baja; un tercer grupo de montaña baja (3), muy próximo ya a la penillanura. Los resultados fueron los siguientes:

GRUPOS 4/1

| | |
|---|---|
| — Altitud/pendiente ... | 0,942 |
| — Altitud/parcelas de —1 ha. ... | 0,771 |
| — Altitud/cultivos leñosos ... | 0,863 |
| — Altitud/encinar y alcornocal ... | —0,933 |
| — Altitud/tierras de labor ... | —0,740 |

GRUPOS 4/2

| | |
|---|---|
| 1.ª ... | 0,801 |
| 2.ª ... | 0,552 |

| | |
|---|---|
| 3.ª | 0,633 |
| 4.ª | —0,767 |
| 5.ª | —0,522 |
| GRUPOS 4/3 | |
| 1.ª | 0,065 |
| 2.ª | 0,096 |
| 3.ª | 0,068 |
| 4.ª | —0,113 |
| 5.ª | —0,056 |

Los resultados son plenamente —creemos— satisfactorios. A partir de ellos, como un punto ya de referencia, se pueden comparar con cualquier otro conjunto o espacio geográfico.

Así pues, podemos concluir que la correlación bidimensional es una operación válida para determinar tanto espacios geográficos homogéneos como diversos, al igual que sus similitudes y diferencias, conscientes— no obstante— de sus limitaciones.

Sería pretencioso el pensar que una operación tan elemental puede competir con los análisis multivariados, tanto en el descubrimiento de las interrelaciones, como de las complejas estructuras que conforman los espacios geográficos.

Pero puede ser de gran utilidad para una aproximación a estas interrelaciones, estructuras, y, en definitiva, a los espacios geográficos. Y, por supuesto, puede sugerir, abundantes hipótesis de trabajo.

De cualquier forma, es una operación delicada, y en todo momento deberá prevalecer la razón y el juicio del investigador.