

4.6.- APLICACION DEL ANALISIS DE COMPONENTES PRINCIPALES EN ECOLOGIA.

(Luis Calabuig, E.)

INTRODUCCION.

En el campo de la Ecología y de las ciencias ambientales es necesaria la utilización de técnicas multivariantes. Una problemática común en los estudios ecológicos radica en la gran cantidad de variables que es necesario considerar, lo que hace prioritario el uso de técnicas que simplifiquen o sintetizen la información, permitiendo, al mismo tiempo, una explicación de los fenómenos o estructuras estudiados.

Entre las técnicas de clasificación y ordenación, quizá la que ha encontrado mayores aplicaciones en Ecología, ha sido la del Análisis en Componentes Principales, por cuanto reúne una serie de características que la hacen muy apropiada a los datos de base obtenidos en la toma de muestras.

Además de otros muchos ejemplos, los datos que recogemos en estudios ecológicos pueden consistir en evaluar distintos lugares de muestra (unidades elementales de muestreo), que definirían las observaciones, con los correspondientes valores de importancia (frecuencia, densidad, cobertura, etc.) de determinados organismos (normalmente de un grupo taxonómico concreto) o una expresión cuantitativa de varios factores ambientales físicos (temperatura, humedad, inclinación, altitud, etc.) o químicos (concentraciones de elementos o productos químicos). Esta visión resumen queda siempre superada por la realidad en cualquier programa de investigación, por lo que recurrir a técnicas de simplificación se nos muestra a los ecólogos

como una necesidad proporcional a la eficiencia de la interpretación o de consecución de objetivos que pretendemos cubrir.

PROBLEMATICA ESPECIFICA.

En algunos casos la eficiencia viene en gran parte definida por el tiempo empleado en la obtención de los datos de base, por lo que cuando el muestreo tiene que ser, por necesidades del espacio o de la variabilidad, muy voluminoso, recurramos a expresiones de tipo cualitativo (siempre que ello sea posible). Esto supone transformaciones posteriores en la matriz de contingencias, como el desdoblamiento en matrices R y Q a partir de presencias, coincidencias o afinidades, resintiéndose en algunos casos el análisis por una pérdida de información.

Pueden expresarse los datos recogidos como una matriz de n filas (observaciones) y p columnas (variables), donde cada elemento representa el valor de importancia de una variable en una observación.

Es muy frecuente que los datos originales necesiten una transformación, incluso cuando partimos de datos cuantitativos, con el objetivo fundamental de normalizar los valores de la distribución. Los procedimientos más aplicados coinciden con los utilizados en otras especializaciones.

Debemos resaltar, sin embargo, la especial problemática que se presenta cuando en la matriz original la frecuencia de ceros es muy elevada. Esto puede plantear, a veces, problemas para los que se han aportado soluciones también muy diversas, pero en algunos casos complicadas y en otros de poco valor para la posterior interpretación de resultados. Un caso típico de presencias elevadas de ceros se nos presenta en el

estudio de gradientes, tanto espaciales como temporales, donde el relevo de especies o de características ambientales es evidente y unidireccional.

También se aplica con cierta frecuencia la transformación logarítmica, para lo que simplemente es necesario añadir una constante, si es que existe algún valor nulo en la matriz original de datos.

Pero quizá en lo referente a los datos de partida, el problema más importante que se nos plantea en Ecología, es que en gran número de ocasiones tenemos más información de partida (fundamentalmente debido al número de variables) de la que el soporte informático disponible puede procesar. Los criterios de simplificación en este caso para reducir las variables hasta una cantidad manejable por la capacidad del ordenador, plantean siempre conflictos de interpretación difíciles de resolver. Por lo general se dejan de tener en cuenta las especies cuya frecuencia relativa de presencias es muy baja y por lo tanto polarizan la información en un sentido muy definido, que puede añadirse después del análisis incluso para reforzar alguna de las tendencias resultantes. Otras veces se eliminan las especies cuyo valor de importancia es relativamente bajo aunque su frecuencia en las observaciones sea elevada. En este caso es más difícil complementar esta información al resultado del análisis con las especies restantes. Un tercer procedimiento de simplificación del número de especies, también muy utilizado, consiste en eliminar previamente las especies presentes en todas las observaciones o de frecuencia relativa muy elevada, en función de que aportan poca información discriminante. Esto sería cierto si el valor de importancia fuera el mismo en todas las observaciones, pero si no es así, eliminarlas puede

