

Alternativa al análisis canónico de correspondencias basada en los métodos Biplot

M.^a J. Fernández Gómez¹,
M.P. Galindo, I. Barrera¹,
J.L. Vicente-Villardón¹,
A. Martín¹

RESUMEN

Los estudios de comunidades ecológicas implican, con frecuencia, el inferir las relaciones existentes entre las especies y las variables ambientales a partir de los datos de composición de la comunidad y de medidas asociadas al hábitat. Es un hecho que la composición de la comunidad cambia a lo largo de *gradientes ambientales* y que las especies tienen dentro de éstos un *óptimo* ambiental y un *rango* concretos. Sin embargo los gradientes ambientales no son directamente observables pero son imprescindibles para poder explicar los reemplazamientos tanto espaciales como temporales de unas especies por otras, cuando varían las condiciones del medio.

En este trabajo, se propone el *HJ-BIPLLOT* como alternativa al Análisis Canónico de Correspondencias (CCA) (ter BRAAK, 1986) siendo también un técnica de análisis multivariante directo del gradiente. El *HJ-BIPLLOT* proporciona mejores resultados que el CCA ya que obtiene la misma calidad de representación para filas y columnas. Se realiza una aplicación a datos reales, y se determinan los *óptimos* de las especies respecto de las variables analizadas y el *ranking* de las mismas en relación al gradiente para dichas variables.

Palabras clave: Análisis Multivariante Directo del Gradiente, Análisis Canónico de Correspondencias. *HJ-BIPLLOT*.

SUMMARY

In this paper, we performe a gradient analysis by using the *HJ-BIPLLOT* as an alternative technique to Canonical Correspondence Analysis (CCA) (ter BRAAK, 1986). The *HJ-BIPLLOT* as a technique of Multivariate Direct Gradient Analysis, provides best results than CCA since it provides the same quality of representation for rows and columns, been much higher than those obtained by others. We apply this method to real data, and determine the species' *optima* respect to the analyzed variables and the *ranking* of the species in relation to the gradient for those variables.

Key Words: Multivariate Direct Gradient Analysis. Canonical Correspondence Analysis. *HJ-BIPLLOT*

INTRODUCCION

Ecólogos de todo el mundo están interesados en los estudios de comunidades, lo cual se traduce en un abanico de escuelas, cada una de ellas con una metodología y énfasis particulares (WHITTAKER, 1962, 1978a,b; MAAREL, 1979; WESTHOFF, 1979 etc.). Pero lo que caracteriza a todas ellas es la progresión hacia el incremento de los métodos cuantitativos (GREIG-SMITH, 1964; ORLOCI, 1978), acelerándose en las últimas décadas debido a la expansión y fácil acceso de los computadores, sumado al importante desarrollo -en los últimos años- de los Métodos Multivariantes. Ahora están disponibles métodos multivariantes efectivos y eficientes para el análisis de datos ecológicos y apropiados para cada problema.

Supongamos que un investigador está interesado en estudiar y describir las relaciones existentes entre dos conjuntos de variables consideradas sobre *n* grupos experimentales. En el estudio de comunidades ecológicas este problema es muy frecuente, ya que se traduce en el estudio de las relaciones entre la composición de la comunidad y las medidas asociadas al hábitat, para lo que es necesario disponer de dos conjuntos de datos medidos sobre las mismas unidades taxonómicas (OTUs): datos de abundancia (o datos de cobertura) de las especies en las diferentes OTUs y datos ambientales (medidas de las variables ambientales de interés en las OTUs).

Las soluciones a este problema encontradas en la bibliografía son muchas y de diversa índole:

El Análisis de Correlación Canónica es una técnica basada en la interpretación de los coeficientes de correlación y por lo tanto requiere que las relaciones existentes entre ambos conjuntos de variables sean de tipo lineal (GITTINS, 1985). Generalmente cuando trabajamos con datos ecológicos esto no se verifica ya que la respuesta de una especie a una variable ambiental es -en general- no lineal.

El modo en que influyen las variables ambientales en la composición biótica de los ecosistemas es a menudo explorada de la siguiente forma: se utilizan técnicas de ordenación (como por ejemplo el Análisis de Componentes Principales (PCA), Análisis Factorial de Correspondencias (CA), métodos Biplot, etc.) para resumir y representar los datos de la composición de la comunidad en un diagrama de ordenación, el cual es posteriormente interpretado según la información ambiental de que se dispone mediante regresión múltiple de las variables ambientales sobre los ejes de ordenación. Esta aproximación en dos pasos es lo que se conoce como ANÁLISIS INDIRECTO DEL GRADIENTE en el sentido utilizado por WHITTAKER (1967), ya que la interpretación ambiental es por lo tanto indirecta. Pero conlleva una serie de problemas:

- Cuando las variables ambientales estén fuertemente

¹ Dpto. de Estadística y Matemática Aplicadas. Universidad de Salamanca.

correlacionadas entre sí, o bien el número de variables se aproxima al número de unidades de muestreo, los coeficientes canónicos son inestables por lo que no es posible conocer el aporte particular de cada una de las variables del estudio. Es el denominado problema de COLINEALIDAD. (CARBONELL et al., 1983).

- Si las variables ambientales están pobremente correlacionadas con los primeros ejes de ordenación obtenidos en el primer paso del estudio, podemos llegar a omitir relaciones importantes.

Frente a este tipo de análisis, tenemos las técnicas que se engloban dentro de lo que se denomina ANÁLISIS DIRECTO DEL GRADIENTE. Con ellas se consigue relacionar de modo directo la variación en la composición de la comunidad con la variación ambiental (GAUCH, 1982). Los métodos más simples de análisis directo, implican representar cada valor de la abundancia de las especies frente a valores de una variable ambiental (un buen ejemplo es el proporcionado por TIVI en 1982), o dibujando gráficos para cada especie en un espacio de dos variables ambientales (BURNETT, 1964; WHITTAKER, 1967). Con estos métodos simples se puede visualizar fácilmente la relación entre muchas especies y una o dos variables. Sin embargo la experiencia nos dice que son muchas las variables ambientales condicionantes de las especies, y por otra parte se puede desear analizar su efecto de manera conjunta. Con este propósito se puede utilizar el análisis de regresión múltiple, sin embargo, a pesar de algunas aplicaciones acertadas, por ejemplo YARRANTON (1970), AUSTIN (1971) y FORSYTHE & LOUCKS (1972), la regresión múltiple ordinaria nunca ha llegado a ser popular en estudios ecológicos. Las razones de ello incluyen:

-El análisis de regresión puede llevarnos a un esfuerzo desmesurado ya que cada especie requiere un análisis por separado.

-En ningún caso los datos de vegetación satisfacen la hipótesis de que la distribución de los errores sea normal, ya que a menudo son cualitativos, o cuando son cuantitativos contienen muchos valores cero para lugares en los cuales la especie está ausente.

-La abundancia de las especies tiene a menudo un único máximo, función de las variables ambientales con lo cual las relaciones entre especies y variables ambientales generalmente no son lineales.

-Las variables ambientales están a menudo altamente correlacionadas entre sí y así puede ser imposible separar sus efectos independientes (problema de colinealidad (CARBONELL et al, 1983)).

El Modelo Lineal Generalizado (ver por ejemplo AUSTIN et al, 1984; ter BRAAK & LOOMAN; 1986) da una solución a los problemas de falta de normalidad en la distribución de los errores y falta de linealidad en la rela-

ción entre especies y variables ambientales, pero sin embargo permanecen los problemas de que cada especie requiere un análisis individual y el de la colinealidad de difícil solución. Además, cuando el número de variables ambientales que influyan sea mayor que dos o tres, será difícil conseguir resultados para varias especies conjuntamente con el objetivo de obtener un gráfico general que resume las relaciones entre especies y variables ambientales.

Es pues necesario un método simple para analizar y visualizar las relaciones existentes entre ambos conjuntos de datos.

Así ter BRAAK propone en 1986 el Análisis Canónico de Correspondencias (CCA) como método de Análisis Multivariante Directo del Gradiente. Se trata de una extensión del Análisis de Correspondencias, una popular técnica de ordenación que extrae ejes continuos de variación a partir de los datos de abundancia u ocurrencia. La matriz utilizada como input en el CCA no es sólo la matriz de especies, lugares sino también una segunda matriz de factores ambientales, lugares. Mediante la utilización de análisis multivariante y técnicas particulares de regresión múltiple, conjuntamente con varias formas de análisis de correspondencias, se obtiene una ordenación integrada de las especies y los datos ambientales asociados. La manera exacta en la que trabaja el CCA es compleja y está explicada con todo detalle en ter BRAAK (1986; 1987).

El autor afirma que esta técnica es una extensión del análisis de correspondencias (alias "medias recíprocas" (RA) HILL (1973)) en la que los ejes de ordenación se eligen en función de las variables ambientales conocidas, imponiendo la restricción adicional de que los ejes deben ser combinaciones lineales de las variables ambientales (*Ordenación restringida*, ter BRAAK, 1988a). De este modo la variación de la comunidad se relaciona directamente con la variación ambiental. Los ejes pueden encontrarse mediante una extensión del algoritmo del PCA (GOURLAY & WATSON, 1973) con un paso adicional en el que se introduce la información procedente de la matriz de datos ambientales, de modo que la regresión no se realiza en un segundo paso -después de la ordenación (como ocurría en el análisis indirecto del gradiente)- sino dentro del propio algoritmo de cálculo de los ejes. Este algoritmo puede estudiarse por ejemplo en JONGMAN & col (1987).

El CCA se convierte en la técnica que selecciona la combinación lineal de las variables ambientales que maximiza la dispersión de las especies (ter BRAAK, 1987). Mientras que el CA es susceptible a lugares "pobres en especies" que contienen especies raras, de modo que estos se colocan como lugares aberrantes en los extremos finales del primer eje de ordenación (GAUCH, 1982) relacionan-

do la mayor variación en los datos con los últimos ejes, el CCA no se ve afectado por este problema, de modo que los lugares que son aberrantes en cuanto a la composición de especies no lo son tanto en términos de las variables ambientales.

GALINDO et al. proponen en 1989 una alternativa al CCA, basada en la utilización del método HJ-BIPLLOT como técnica de Análisis Multivariante Directo del

Gradiente, y que es posteriormente desarrollada por FERNANDEZ-GOMEZ (1990) donde se demuestran las ventajas del método en relación con el CCA y con otros análisis.

Para facilitar el entendimiento, clarificar las ventajas y destacar la importancia de dicha técnica en las aplicaciones biológicas se aplicarán estos conceptos a un ejemplo concreto.

- | | |
|--|---|
| 1 <i>Agrostis Castellana</i> B. et R. | 44 <i>Gaudinia fragilis</i> (L.) P.B. |
| 2 <i>Trifolium striatum</i> L. | 45 <i>Agrostis salmantica</i> (lag.) Kunth. |
| 3 <i>Moenchia erecta</i> (L) Gaertn. | 46 <i>Hypochaeris glabra</i> L. |
| 4 <i>Vulpia bromoides</i> (L) S.F. Gray | 47 <i>Galium verum</i> L. |
| 5 <i>Hypochaeris radicata</i> L. | 48 <i>Trifolium parviflorum</i> Ehrh. |
| 6 <i>Bromus mollis</i> L. | 49 <i>Daucus carota</i> L. |
| 7 <i>Festuca rubra</i> L. | 50 <i>Trifolium scabrum</i> L. |
| 8 <i>Plantago lanceolata</i> L. | 51 <i>Cerastium diffusum</i> Pers. |
| 9 <i>Cerastium glomeratum</i> Thuill | 52 <i>Lotus corniculatus</i> L. |
| 10 <i>Thrinicia hispida</i> Roth | 53 <i>Holcus setiglomis</i> Bss. Reut. |
| 11 <i>Trifolium glomeratum</i> L. | 54 <i>Vulpia myuros</i> (L.) Gmel. |
| 12 <i>Trifolium strictum</i> L. | 55 <i>Hordeum secalinum</i> Schreb. |
| 13 <i>Anthoxanthum aristatum</i> Boiss. | 56 <i>Poa trivialis</i> L. <i>ulbosum</i> (Willd.) Koch. |
| 14 <i>Trifolium dubium</i> Sibth. | 57 <i>Trifolium resupinatum</i> L. |
| 15 <i>Thrinicia hirta</i> Roth. | 58 <i>Linum gallicum</i> L. |
| 16 <i>Aira caryophyllea</i> L. | 59 <i>Arrhenatherum elatius</i> (L.) Var. <i>bulbosum</i> (Willd) Koch. |
| 17 <i>Poa bulbosa</i> L. | 60 <i>Brassica barrelieri</i> (L.) Janka. |
| 18 <i>Crepis capillaris</i> (L.) Wallr. | 61 <i>Trifolium fragiferum</i> L. |
| 19 <i>Trifolium micranthum</i> Viv. | 62 <i>Merendera montana</i> (L.) Lange. |
| 20 <i>Trifolium subterraneum</i> L. | 63 <i>Filago germanica</i> L. subsp. <i>spatulata</i> (Press) Rouy. |
| 21 <i>Parentucellia latifolia</i> (L.) Car. | 64 <i>Myosotis discolor</i> Pers. |
| 22 <i>Aphanes microcarpa</i> (B. et R.) Rothm. | 65 <i>Scleranthus annuus</i> L. |
| 23 <i>Convolvulus arvensis</i> L. | 66 <i>Anthemis nobilis</i> L. |
| 24 <i>Phleum nodosum</i> L. | 67 <i>Periballia laevis</i> (Brot.) Asch. et Graebnt. |
| 25 <i>Eryngium campestre</i> L. | 68 <i>Spergula arvensis</i> L. |
| 26 <i>Linum Bienne</i> Miller. | 69 <i>Teesdalia coronopifolia</i> (Berg.) Tell. |
| 27 <i>Tuberaria guttata</i> (L. | 70 <i>Alopecurus pratensis</i> L. |
| 28 <i>Ornithopus compressus</i> L. | 71 <i>Prunella lacinicata</i> (L.) L. |
| 29 <i>Ornithopus perpusillus</i> L. | 72 <i>Anthemis mixta</i> L. |
| 30 <i>Cynodon dactylon</i> (L.) Pers. | 73 <i>Elymus caput-medusae</i> L. |
| 31 <i>Anthemis arvensis</i> L. | 74 <i>Antyllis lotoides</i> L. |
| 32 <i>Trifolium campestre</i> Schereb. | 75 <i>Vulpia ciliata</i> Link. |
| 33 <i>Galium divaricatum</i> Lam. | 76 <i>Parentucellia viscosa</i> (L.) Caruel. |
| 34 <i>Carex chaetophylla</i> Steud. | 77 <i>Rumex angiocarpus</i> Murb. |
| 35 <i>Trifolium arvense</i> L. | 78 <i>Carlina racemosa</i> L. |
| 36 <i>Veronica arvensis</i> L. | 79 <i>Rhinanthus minor</i> L. |
| 37 <i>Narcissus bulbocodium</i> L. | 80 <i>Ranunculus bulbosus</i> L. |
| 38 <i>Silene gallica</i> L. | 81 <i>Trifolium pratense</i> L. |
| 39 <i>Euphorbia exigua</i> L. | 82 <i>Conopodium capillifolium</i> Boiss. |
| 40 <i>Bellis perennis</i> L. | 83 <i>Cynosurus cristatus</i> L. |
| 41 <i>Centaurea ornata</i> Willd. | 84 <i>Trifolium repens</i> L. |
| 42 <i>Erodium ciconium</i> (L.) L'Hér. | 85 <i>Tolpis barbata</i> (L.) Gaertn |
| 43 <i>Filago minima</i> Pers. | 86 <i>Petrorrhagia prolifera</i> (L.) Ball. and Heywood |

Tabla 1. Relación de especies.

MATERIAL

Los datos utilizados en el trabajo han sido tomados de LUIS (1976). En el cual se recogieron datos de cobertura (tanto por ciento de terreno cubierto por cada especie en cada comunidad), y datos sobre las características de los suelos donde estas comunidades se asientan (variables edáficas). En total se encontraron 86 especies diferentes (Ver tabla 1).

Las variables de suelo analizadas son las clásicas en este tipo de estudios *pH* (1), % de Nitrógeno (2), Relación Carbono-Nitrógeno (C/N) (3), % de Materia Orgánica (4), CaO (5), P_2O_5 (6), K_2O (7) y fracciones granulométricas: Arena gruesa (8), %Arena Fina (9), % Limo (10) y %Arcilla (11).

También se disponía de datos de dos variables climáticas: Precipitación (12) y Temperatura (13) las cuales se introducen en el estudio conscientes de la falta de precisión, por la forma en la que estaban medidas (Ver LUIS, 1976).

La estructura de los datos es, pues, la siguiente:

- Disponemos de una matriz $Y_{20 \times 86}$ de datos de cobertura, es decir, tanto por ciento de terreno cubierto por cada especie en cada lugar.

- Datos ambientales en forma matricial: $Z_{20 \times 12}$; siendo 12 las variables ambientales, edáficas y climáticas anteriormente señaladas.

En el estudio se pretende determinar los "OPTIMOS" de las especies respecto de las variables ambientales estudiadas y establecer el "RANKING" de las especies en relación al gradiente para dichas variables. De este modo podremos conocer la relación existente entre la composición de las comunidades y las variables ambientales estudiadas.

METODOLOGIA ESTADISTICA

Se desecha de entrada la aplicación del Análisis de Correlación Canónica ya que la respuesta de las especies frente a factores ambientales no sigue un modelo lineal, y sería incorrecto utilizar esta técnica basada en la interpretación de los coeficientes de correlación.

Cualquier técnica de ordenación (PCA, CA, métodos BIPLLOT) puede ser útil para obtener el patrón dominante en la composición de la comunidad, y después interpretar dicho patrón a la vista de la información ambiental disponible para realizar un Análisis Indirecto del Gradiente (IGA) (WHITTAKER, 1967).

Esta opción conlleva una serie de problemas obvios en este caso concreto ya que las variables de suelo están fuertemente correlacionadas.

Optamos pues por un Análisis Directo del Gradiente, en el que los datos de las especies se relacionan directamente con las variables ambientales.

ter BRAAK (1986) demuestra, que llevar a cabo un CCA, es equivalente a realizar un GH'-BIPLLOT (GABRIEL, 1971) de la matriz de estimadores de los óptimos para las especies en relación a las variables estudiadas. En esta representación BIPLLOT, las variables están bien representadas, es decir, su calidad de representación (CLR) es alta pero la CLR para las filas (especies) es mucho más baja (GABRIEL & ODOROFF, 1990).

Para interpretar el gradiente para una variable en el gráfico de ordenación resultante, es necesario proyectar el punto que representa a una especie sobre el vector que representa a una variable (JONGMAN & col. 1987), por lo que si la posición del punto-especie es sólo aparente, el "ranking" de las especies en relación al gradiente para dicha variable puede no ser fiable. Sin embargo, el fácil acceso al programa computacional CANOCO (ter BRAAK, 1988b,c; 1990) hace que la mayor parte de los usuarios utilicen esta técnica sin tener en cuenta cuáles son sus limitaciones.

Se hace necesaria la utilización de una técnica que permita una representación simultánea, y en la que tanto las especies como las variables ambientales vengan bien representadas.

La solución propuesta por GALINDO et al. (1989) se basa en la utilización del HJ-BIPLLOT (GALINDO, 1985) como técnica de ordenación para encontrar los patterns dominantes de variación en la composición de la comunidad, siendo posible representar en el mismo diagrama de ordenación las variables ambientales. Se trata de una técnica de representación simultánea en sentido estricto.

Las ventajas están en el hecho de que el HJ-BIPLLOT obtiene la misma calidad de representación tanto para las especies como para las variables ambientales, siendo mucho más alta que la conseguida con cualquier otra técnica de representación.

En el problema que nos ocupa, disponemos de dos matrices de datos $Y_{(n \times m)} = [y_{ik}]$ y $Z_{(n \times p)} = [z_{ij}]$ donde:

y_{ik} = abundancia de la especie k-ésima en el lugar i-ésimo.

z_{ij} = valor de la variable ambiental j-ésima en el lugar i-ésimo.

Evidentemente no es posible observar el valor de la variable ambiental j-ésima en la especie k-ésima, sin embargo, no es menos evidente, que la respuesta de la especie está claramente relacionada con la variable y que existe un óptimo ambiental para el que se maximiza la abundancia de la especie.

Así pues, se utilizará como matriz input en el análisis HJ-BIPLLOT la matriz de medias ponderadas $W_{(m \times p)}$ formada por los estimadores de los óptimos para las especies en

relación a las variables ambientales. Los elementos de dicha matriz se calculan:

$$w_{jk} = \sum_i y_{ik} z_{ij} / y_{.k}$$

siendo $y_{.k}$ la abundancia total de la especie k -ésima.

Expresado en forma matricial:

$$W = S_{jj}^{-1} Y' Z$$

Donde S_{jj} es una matriz diagonal cuyos elementos son los $y_{.k}$

Se pretende encontrar una aproximación mínimo cuadrática de W en un diagrama de ordenación.

Sin embargo, la media ponderada para especies poco abundantes es imprecisa (ver BRAAK, 1986), por lo que será necesario hacer una corrección en este sentido, dando a las especies pesos que son proporcionales al total de las especies contenido en $S_{jj}^{1/2}$. Así mismo, las columnas de esta matriz no son comparables ya que vienen medidas en diferentes unidades. Para hacer el método invariante por el cambio de escala se usa $S_{22}^{-1/2}$ como pesos para las variables ambientales, donde S_{22} es la matriz de covarianzas entre las variables ambientales.

Por ello, la matriz ponderada resultante será:

$$\mathcal{W} = S_{jj}^{1/2} W S_{22}^{-1/2}$$

La deseada aproximación mínimo cuadrática de \mathcal{W} ahora proviene de la descomposición en valores singulares de la matriz \mathcal{W}

$$\mathcal{W} = U \Sigma V'$$

donde: U es la matriz de vectores propios de $\mathcal{W}\mathcal{W}'$

V es la matriz de vectores propios de $\mathcal{W}'\mathcal{W}$

Σ matriz diagonal cuyos elementos son los valores singulares.

La matriz \mathcal{W} puede representarse en el subespacio E2 mediante una HJ-Bigeometría: (GALINDO, 1986; VICENTE-VILLARDON, 1992)

$$H = U \Sigma$$

$$J = V \Sigma$$

Las matrices \mathcal{W} y \mathcal{W}' están relacionadas, y su descomposición en valores singulares también está relacionada:

$$W = S_{jj}^{-1/2} \mathcal{W} S_{22}^{1/2} = S_{jj}^{-1/2} U \Sigma V' S_{22}^{1/2}$$

de forma que las coordenadas para las especies y las variables ambientales en la descomposición de W puede escribirse como:

$$A(\text{especies}) = S_{jj}^{-1/2} U \Sigma$$

$$B(\text{variables}) = S_{22}^{1/2} V \Sigma$$

Utilizando las relaciones que ligan U con V :

$$U = \mathcal{W} V \Sigma^{-1}$$

$$V = \mathcal{W} U \Sigma^{-1}$$

$$A(\text{especies}) = S_{jj}^{-1/2} U \Sigma = S_{jj}^{-1/2} \mathcal{W} V \Sigma \Sigma^{-1} = S_{jj}^{-1/2} \mathcal{W} V = S_{jj}^{-1/2} S_{jj}^{1/2} W S_{22}^{-1/2} V = W S_{22}^{-1/2} V$$

$$B(\text{variables}) = S_{22}^{1/2} V \Sigma = S_{22}^{1/2} \mathcal{W} U \Sigma \Sigma^{-1} = S_{22}^{1/2} S_{22}^{-1/2} W' S_{jj}^{1/2} U = W' S_{jj}^{1/2} U$$

Entonces

$$A(\text{especies}) = W S_{22}^{-1/2} S_{22}^{-1/2} S_{22}^{1/2} V \Sigma \Sigma^{-1} = W S_{22}^{-1} B \Sigma^{-1}$$

$$B(\text{variables}) = W' S_{jj}^{1/2} S_{jj}^{1/2} S_{jj}^{-1/2} U \Sigma \Sigma^{-1} = W' S_{jj}^{1/2} A \Sigma^{-1}$$

Las ecuaciones anteriores nos permiten representar las coordenadas de las filas y de las columnas con referencia al mismo sistema de ejes cartesianos, es decir, los espacios de representación de las filas y de las columnas pueden superponerse en orden a obtener una representación conjunta ya que los valores propios para el análisis de las nubes de puntos permiten identificar los ejes factoriales correspondientes al mismo valor propio y representar ambas nubes en el mismo diagrama.

El HJ-BIPLLOT preserva la métrica de las columnas y la métrica de las filas: las varianzas-covarianzas, las distancias estandarizadas entre la i -ésima y la i' -ésima filas (especies) y las distancias entre las j -ésima y j' -ésima columnas (variables) están bien representadas (GALINDO y CUADRAS, 1986).

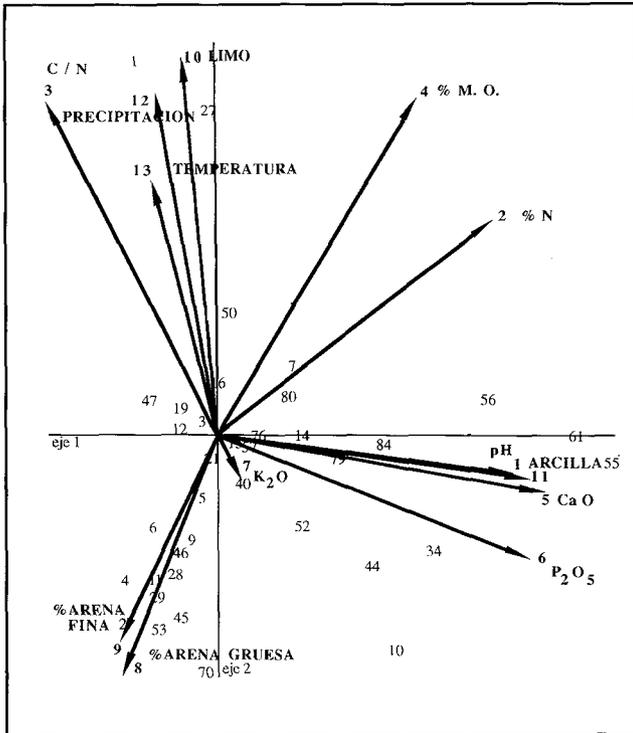
RESULTADOS Y DISCUSION

El Análisis Directo del Gradiente, basado en el análisis HJ-BIPLLOT de la matriz de estimadores de los óptimos para las especies en estudio, proporcionó el diagrama de ordenación que puede verse en la gráfica 1.

El primer eje principal absorbe una inercia del 68.14% y el segundo eje un 10.86%, de modo que la interpretación se basará en el primer plano principal.

La bondad de ajuste tanto para las especies como para las variables es muy alta: supera el 97%.

Las variables vienen representadas por vectores, que se construyen uniendo el punto que representa a la variable



Gráfica 1. Proyección en el plano de máxima inercia de los puntos que representan a las especies y a las variables según el HJ-BIBLOT de la matriz X (especies x variables).

con el centro de gravedad de la hipernube proyectada sobre el subespacio de máxima inercia. La posición de la punta de flecha marcará el *extremo superior del gradiente para dicha variable*.

La longitud del vector es una medida de la variabilidad de los estimadores de los óptimos a lo largo de la variable ambiental.

Las especies vienen representadas con puntos.

Es muy importante analizar si la posición de los puntos que representan a las variables y/o a las especies es interpretable o sólo aparente. Las especies cuya calidad de representación es más alta son la 1, 2, 10, 11, 28, 29, 34, 44, 45, 46, 50, 52, 53, 55, 56, 61, 80 y 84. Para detalle ver la gráfica 1.

Las variables características de eje I son: *pH*, *CaO*, *%N*, *Arcilla* y *P₂O₅* y las características de eje II: *Limo*, *Temperatura* y *Precipitación*.

La variable relación C/N no es característica de ninguno de los dos ejes pero sí del plano. La variable *K₂O* se encuentra mal representada en el primer plano factorial por lo que se omitirán relaciones al respecto.

Hay que destacar que el pequeño ángulo que forman entre sí las variables *CaO*, *pH*, *Arcilla* y *P₂O₅* indica una

alta correlación entre dichas variables. Lo mismo sucede entre las variables *%N* y *%M.O.*

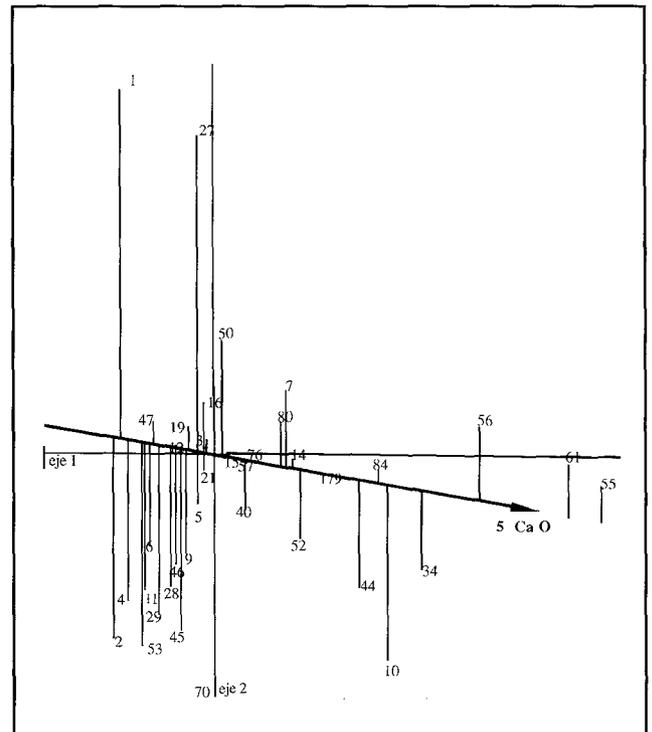
Desde cada punto que representa a una especie, debemos trazar la perpendicular al vector que determina a una variable. Los puntos de corte de dichas perpendiculares con ese vector nos indican de forma aproximada el valor relativo de la media ponderada de cada especie con respecto a la variable ambiental.

Los puntos que representan a las especies, cuya proyección sobre el vector variable ambiental queden más próximos a la punta de la flecha nos indican que tienen óptimos más altos con respecto a dicha variable.

En general el "ranking" aproximado de las medias ponderadas para una variable ambiental particular se deduce del orden de las proyecciones sobre el correspondiente vector. Esta propiedad evidencia la importancia de utilizar una técnica de representación simultánea en sentido estricto.

En la gráfica 2 podemos ver el "ranking" de las especies en relación al gradiente para la variable *CaO*, con el fin de clarificar los resultados. En él podemos ver como las especies 55, 61, 56, 10, 84, 44, 79 y 52 marcan el extremo superior del gradiente para dicha variable.

Recordemos el pequeño ángulo que formaban las variables *pH*, *CaO*, *Arcilla* y *P₂O₅* nos indicaba una alta



Gráfica 2. Diagrama de ordenación para el Análisis del Gradiente de la variable *CaO*.

Especie	pH	%N	C/N	%M.O.	CaO	P ₂ O ₅	K ₂ O	%Ar. Gruesa	% Ar. Fina	% Limo	% Arcilla
<i>Hordeum secalinum</i>	7,4	,57	8,82	8	838,11	7,24	23,5	9,8	13,97	21,57	45,99
<i>Trifolium fragiferum</i>	7,4	,6	8,81	9	869,51	7,44	24,46	8,08	13,25	22,26	47,35
<i>Gaudinia fragilis</i>	6,91	,3	9,79	4,91	418,76	4,3	13,08	24,47	22,42	19,28	28,13
<i>Poa trivialis</i>	7,07	,56	9	8,56	757,63	6,63	22,88	7,4	13,77	25,96	42,39
<i>Carex chaetophylla</i>	6,74	,32	9,73	5,12	444,12	4,76	24,18	20,91	22,83	21,23	29,66
<i>Thrinicia hispida</i>	6,98	,27	9,43	4,38	418,46	4,42	12,67	29,27	22,61	15,83	28,04
<i>Trifolium repens</i>	7,34	,54	8,9	8,27	793,63	6,93	22,28	11,01	15,04	21,62	44,08
<i>Rhinanthus minor</i>	6,42	,28	10,5	4,84	350,87	3,17	11,15	14,95	19,84	30,15	30,3

Tabla 2. Estimadores de los óptimos de las especies características de suelos ricos.

covariación entre ellas, por lo que aquello que decimos respecto de la variable CaO se puede hacer extensible a las anteriormente indicadas.

Los diagramas de ordenación para el análisis del gradiente con respecto al resto de las variables se realizaría de modo análogo al anterior.

Si analizamos el gradiente para la variable % de Arena Gruesa podemos observar como las especies que marcaban el extremo superior para el pH, CaO, Arcilla y P₂O₅, marcan ahora el extremo inferior del gradiente para esta variable. La información para el % de Arena Fina es análoga, dada la fuerte correlación con la anterior.

En relación al gradiente para la variable % de Materia Orgánica puede verse como esas especies marcan también el extremo superior del gradiente de esta variable. Información análoga se encuentra en el gradiente para la variable relación Carbono Nitrógeno.

En cuanto a la fracción Limo estas especies forman parte del conjunto del extremo inferior del gradiente, aunque no guardan el mismo orden.

A la vista de los resultados encontrados podemos afirmar que ese grupo de especies es característico de suelos ricos. En la tabla 2 pueden leerse los estimadores de los óptimos para dichas especies. Para alguna de estas especies disponemos de información, recogida en la literatura, la cual es compatible con la encontrada en el estudio, salvo en algún caso que señalaremos en la descripción.

Hordeum secalinum, aparece citada como una especie característica de prados semiagostantes; ciertamente aparece con frecuencia en ambientes con cierto grado de salinidad como consecuencia de la acumulación de sales por efecto de la fuerte evapotranspiración. *Trifolium fragiferum* y *Gaudinia fragilis*, también características de prados semiagostantes, es frecuente encontrarlas en praderas semiagostantes aprovechadas para pastoreo; (LUIS, op. cit.) afirma que suelen asentarse sobre suelos profundos ricos en fracciones granulométricas finas donde perdura la

humedad edáfica, sin embargo en nuestro estudio aparecen en el extremo inferior del gradiente para arena fina. La disparidad de resultados no puede ser achacada a una mala calidad de representación de las especies puesto que son de las mejor representadas. *Poa trivialis*, también característica de prados semiagostantes, es típica de comunidades asentadas sobre suelos profundos inundados durante el invierno y la primavera. Suele encontrarse en praderas semiagostantes que bordean arroyos de cauce discontinuo. *Carex chaetophylla*, aparece en la literatura como especie característica de majadales es decir pastizales que soportan carga ganadera. *Thrinicia hispida*, es señalada como especie de vallicares y majadales, de características similares a las de los prados semiagostantes aunque algo menos húmedos que los clásicos. *Trifolium repens* y *Rhinanthus minor* son referidas como especies de prados semiagostantes, aunque la segunda los prefiere menos húmedos y *Lotus corniculatus* como especie de pastizales más evolucionados. Todas ellas son, por tanto, especies representantes de comunidades evolucionadas o maduras, lo cual concuerda con nuestros resultados.

Las especies 2, 3, 4, 6, 29, 28, 46 marcan el extremo inferior del gradiente de las variables CaO, Arcilla, P₂O₅, pH y Limo. Podríamos decir, por tanto, que son especies típicas de los suelos más pobres. Todas estas especies aparecen en el extremo superior del gradiente para las variables Arena gruesa y Arena fina. Los estimadores de sus óptimos en relación a las variables ambientales estudiadas puede verse en la tabla 3.

Con el fin de contrastar nuestros resultados con los encontrados en la literatura señalaremos que *Trifolium striatum* aparece como especie típica de vallicares normales, *Moenchia erecta* aparece en cualquier lugar que permita algo de materia orgánica para subsistir, se asienta sobre suelos poco profundos, ricos en fracciones granulométricas gruesas donde no perdura la humedad edáfica. En este caso, nuestros resultados en la relación a las fraccio-

Especie	pH	%N	C/N	%M.O.	CaO	P2O5	K2O	%Ar. Gruesa	% Ar. Fina	% Limo	% Arcilla
Trifolium striatum	5,17	,17	10,86	3,17	95,83	1,41	15,77	32,58	24,06	22,94	16,57
Moenchia erecta	5,34	,22	11,35	4,19	138,41	1,54	14,29	19,41	25,63	30,73	19,72
Vulpia bromoides	5,23	,2	10,8	3,76	106,68	1,45	20,31	19,82	30,76	27,86	17,74
Bromus mollis	5,3	,19	11,23	3,74	113,08	1,18	19,8	12,76	40,95	24,16	18,68
Ornithopus perpusillus	4,99	,14	10,51	2,47	52,57	1,37	9,49	41,78	23,12	16,99	13,37
Ornithopus compressus	4,82	,12	10,75	2,24	41,2	1,36	8,57	46,09	23,85	14,36	11,93
Hypochaeris glabra	4,81	,14	11,05	2,67	41,53	1,51	10,12	44,53	24,38	14,88	12,05

Tabla 3. Estimadores de los óptimos de las especies características de suelos pobres.

nes granulométricas coinciden. *Vulpia bromoides*, y *Bromus mollis* especies típicas de vallicares de siega, con bastante humedad, suelo profundo y muy estabilizados; este resultado es contrario al encontrado al hacer la lectura del gradiente, sin embargo en este caso hay una clara explicación y es que estas especies está muy mal representadas en este plano por eso su posición es sólo aparente y no debe ser leída en ninguno de los gráficos. *Ornithopus perpusillus*, *Ornithopus compressus* e *Hypochaeris glabra*, son típicas de comunidades de efímeras, es decir, comunidades que se asientan sobre suelos muy pobres, poco profundos, arenosos y con poca capacidad de retención, características todas ellas que concuerdan con nuestros resultados.

AGRADECIMIENTOS

Dada la enorme dificultad con la que a veces se encuentra un equipo de trabajo de estadísticos a la hora de acceder a datos que pongan de manifiesto la importancia de sus aportaciones teóricas a nivel práctico, deseamos mostrar nuestro más sincero agradecimiento al Doctor D. Estanislao Luis Calabuig por habernos permitido utilizar los datos de su Tesis Doctoral para aplicar la metodología desarrollada en este trabajo.

BIBLIOGRAFIA

AUSTIN, M.P., (1971). Role of regression analysis in plant ecology. *Proc. Ecol. Soc. Austr.*, 6: 63-75.
 AUSTIN, M.P. CUNNINGHAM, R.B. & FLEMING, P.M., (1984). New approaches to direct gradient analysis using environmental scalars and statistical curve-fitting procedure'. *Vegetatio*, 55: 11-27.
 BURNETT, J.D. (ed.), (1964). *The Vegetation of Scotland*. Oliver and Boyd, Edinburgh.
 CARBONELL, E.; DENIS, J.B.; CALVO, R.; GONZALEZ, F. y PRUÉEONOSA, V., (1983). *Regresión Lineal: Un enfoque conceptual y práctico*. I.N.I.A. España.
 FERNANDEZ-GOMEZ, M.J., (1990). *Contribuciones al Análisis Multivariante Directo del Gradiente basadas en los Métodos Biplot*. Tesina de Licenciatura. Univ. de Salamanca.
 FORSYTHE W.L. & LOUCKS O.L., (1972). A transformation for species response to habitat factors. *Ecology* 53: 1112-1119.

GABRIEL, K.R., (1971). The biplot graphic display of matrices with applications to principal component analysis'. *Biometrika*. 58: 453-467.
 GABRIEL, K.R. & ODOROFF, CH. L., (1990). Biplots in biomedical research. *Statistics in Medicine* Vol. 9: 469-485.
 GALINDO, M.P., (1985). Contribuciones a la representación simultánea de datos multidimensionales. *Ser. Resúm. Tesis Doct.* T.D. 395/1985. Pgs. 1-38.
 GALINDO, M.P., (1986). Una alternativa de representación simultánea: HJ-BIPLLOT. *Questiio*. Vol. 10, Nº 1: 13-23.
 GALINDO, M.P. y CUADRAS, C.M., (1986). Una extensión del método Biplot y su relación con otras técnicas. *Publicaciones de Bioestadística y Biomatemática*. Un. Barcelona, nº 17.
 GALINDO, M.P.; VICENTE-VILLARDON, J.L. y BARRERA, I., (1989). La representación "HJ-BIPLLOT" como técnica de Análisis Multivariante Directo del Gradiente. *Actas de la II Conferencia Nacional de Biometría*. Segovia.
 GAUCH, H.G., (1982). *Multivariate Analysis in Community Ecology*. Cambridge University Press.
 GITTINS, R., (1985). *Canonical Analysis. A review with applications in Ecology*. Springer-Verlag. Berlin, Germany.
 GOURLAY, A.R. & WATSON, G.A., (1973) *Computational Methods for Matrix Eigen Problems*. John Wiley. New York.
 GREIG-SMITH, P., (1964). *Quantitative Plant Ecology*, 2d. ed. Butterworths. London.
 HILL, M.O., (1973). Reciprocal averaging: An eigenvector methods of ordination. *Journal of Ecology*. 61: 237-249.
 JONGMAN, R.H.G.; ter BRAAK, C.J.F. & TONGEREN, F. R. van, (1987). *Data analysis in community and landscape ecology*. Pudoc.Wageningen.
 LUIS, E., (1976). *Ecosistemas de pastizal: Estudio de la vegetación mediante técnicas de Análisis Factorial. Crecimiento primario*. Tesis Doctoral. Univ. de Salamanca.
 MAAREL, E. van der, (1979). Multivariate methods in phytosociology, with reference to The Netherland'. *In The Study of vegetation*, ed. M.J.A. Werger, pp. 161-225. Junk. The Hague.
 ORLOCI, L., (1978). *Multivariate Analysis in Vegetation Research*. 2d ed. Junk. The Hague.
 ter BRAAK C.J.F., (1986). Canonical Correspondence Analysis: a new eigenvector -technique for multivariate direct gradient analysis. *Ecology*, 67(5): 1167-1179.
 ter BRAAK, C.J.F., (1987). The analysis of vegetation-environment relationships by canonical correspondence analysis. *Vegetatio*, 64: 69-77.
 ter BRAAK, C.J.F., (1988a). Partial canonical correspondence analysis. *In Classification Methods and Related Methods of Data analysis* (Ed. by H.H. Bock) Amsterdam: North-Holland, 551-558.

- ter BRAAK, C.J.F., (1988b). *CANOCO -a FORTRAN Program for Canonical Community Ordination by (Partial) (Detrended) (Canonical) Correspondence Analysis. (Version 2.0)*. TNO Institute of Applied Computer Science, Wageningen.
- ter BRAAK, C.J.F., (1990). *Update Notes: CANOCO version 3.1*. Agricultural Mathematics Group, Wageningen.
- ter BRAAK, C.J.F., (1988c). CANOCO -an extension of DECORANA to analyse species-environment relationships. *Vegetatio*, 75: 159-160.
- ter BRAAK, C.J.F. & LOOMAN, C.W.N.; (1986). Weighted averaging, logistic regression and the Gaussian response model. *Vegetatio*, 65: 3-11.
- TIVI, J. , (1982). *Biogeography: a Study of Plants in the Ecosphere*. 2nd ed., Longman, London.
- VICENTE-VILLARDON, J.L., (1992). *Una alternativa a las técnicas factoriales clásicas basada en una generalización de los métodos BIPLLOT*. Tesis Doctoral. univ. de Salamanca.
- WESTHOFF, V., (1979). Phytosociology in The Netherlands: History, present state, future. In *The Study of Vegetation*, ed. M.J.A. Werger, pp. 81-121. Junk. The Hague.
- WHITTAKER, R.H., (1962). Classification of natural communities. *Botanical Review*, 28: 1-239.
- WHITTAKER, R.H., (1967). Gradient analysis of vegetation. *Biological Reviews of the Cambridge Philosophical society*, 42: 207-264.
- WHITTAKER, R.H. Ed., (1978b). *Ordination of Plant Communities*. Junk. The Hague.
- YARRANTON, G.A., (1970). Towards a mathematical model of limestone pavement vegetation. III. Estimation of the determinants of species frequencies. *Can J. Bot.* 48: 1387-1404.