

## INTERPRETANDO LOS RESULTADOS DE PCA

El resultado de la ordenación utiliza un vocabulario que requiere explicaciones.

- **Inertia** (inercia): en el lenguaje de *vegan*, este es el término general para "variación" en los datos. Este término proviene del mundo del análisis de correspondencia. En PCA, la "inercia" es la suma de las varianzas de las variables (PCA en una matriz de covarianza) o, como en este caso (PCA en una matriz de correlación), la suma de los valores diagonales de la matriz de correlación, es decir, la suma de todas las correlaciones de las variables con ellas mismas, que corresponde al número de variables (11 en este ejemplo).
- **Constrained and unconstrained** (restringido y no restringido): En PCA, el análisis no está restringido, es decir, no está limitado por un conjunto de variables explicativas, y tampoco lo están los resultados.
- **Eigenvalues** (valores propios), simbolizados  $\lambda_j$ : son medidas de la importancia (varianza) de los ejes PCA. Se pueden expresar como **proporciones explicadas**, o proporciones de variación explicadas por los ejes, dividiendo cada valor propio por la "inercia total".
- **Scaling** (escalado): no debe confundirse con el argumento "scale" que exige la estandarización de variables. "Escalado" se refiere a la forma en que se proyectan los resultados de la ordenación en el espacio reducido para la visualización gráfica. No hay una forma única de mostrar objetos y variables de manera óptima en un biplot de PCA, es decir, una gráfica que muestre dos tipos de resultados, aquí los sitios y las variables. Generalmente se utilizan dos tipos principales de escalado. Cada uno de ellos tiene propiedades que deben tenerse en cuenta para la correcta interpretación de los biplots. Aquí damos las características esenciales de cada escalado. Consultar Legendre y Legendre (2012, p. 443–445) para obtener más detalles:
  - **Scaling 1** (escalado 1) = distancia de biplot: los vectores propios se escalan a la longitud de la unidad:

- Las distancias entre los objetos en el biplot son aproximaciones de sus distancias euclidianas en el espacio multidimensional.
  - Los ángulos entre los vectores descriptores no reflejan sus correlaciones.
- **Scaling 2** (escalado 2) = correlación de biplot: cada vector propio se escala a la raíz cuadrada de su valor propio.
- Las distancias entre los objetos en el biplot no son aproximaciones de sus distancias euclidianas en el espacio multidimensional.
  - Los ángulos entre los descriptores en el biplot reflejan sus correlaciones.
- En ambos casos, proyectar un objeto en ángulo recto en un descriptor se aproxima a la posición del objeto a lo largo de ese descriptor.
- Línea inferior: si el interés principal del análisis es interpretar las relaciones entre los objetos, elija la escala 1. Si el interés principal se centra en las relaciones entre los descriptores, elija la escala 2.
- Hay una escala de compromiso, **Scaling 2** (escala 3), también llamada "escala simétrica", que consiste en escalar las puntuaciones de sitio y de especie por las raíces cuadradas de los valores propios. Se supone que esta escala permite una representación simultánea de sitios y puntajes sin enfatizar uno u otro punto de vista. Este escalamiento de compromiso no tiene reglas de interpretación claras y, por lo tanto, no lo discutiremos más en este curso.
- **Species scores** (puntuaciones de especies): coordenadas de las puntas de flecha de las variables. Por razones históricas, las variables de respuesta siempre se llaman "species" en vegan, sin importar lo que representen, porque vegan contiene software para el análisis de la vegetación.

- **Site scores** (puntuaciones del sitio): coordenadas de los sitios en el diagrama de ordenación. Los objetos siempre se llaman "Sites" en archivos de salida vegan.