



1. En este problema se utilizaron los datos de 21 universidades y los delitos (10) donde se vieron involucrados los estudiantes de dichas universidades en el año 2008.

El primer paso consistió en calcular los crímenes per cápita por universidad, dividiendo los diferentes crímenes entre la matricula de estudiantes. Posteriormente se estandarizaron los datos por tipo de crimen y se obtuvo una matriz de distancias (euclidiana), es decir resulta en una matriz de tamaño 21x21.

Por ultimo con la anterior matriz de distancias se aplicó el algoritmo SMACOF y se evaluó el número de dimensiones y mediante el grafico de la Figura 1.1 a), se determinó 5 dimensiones, ya que es el punto de cambio de pendiente y se realizó el diagrama de Shepard para evaluar el modelo, observe la Figura 1.1 b) y note que el ajuste es muy bueno, prácticamente todos los puntos caen sobre la línea recta y esto corresponde con el Stress-1 de 0.006 que es perfecto.

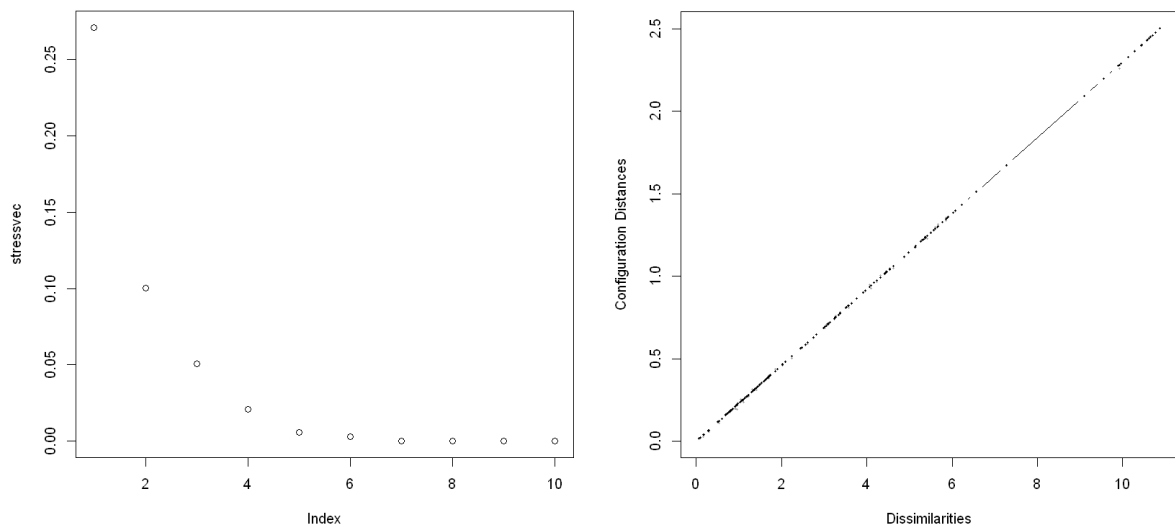


Figura 1.1: a) Stress-1 vs dimensionalidad. b) Diagrama de Shepard para dimensión 5.

La solución se muestra en la Figura 1.2, en la primera dimensión para el eje horizontal observando los datos, se notó que es respecto a la violencia, es decir, mientras mas a la derecha se encuentre una universidad mas violentos son los crímenes, en este sentido la universidad Benedict es la que parece mucho más violenta (también podría ser un dato atípico) y también la SCState. Por otro lado, hasta el extremo izquierdo se nota un

2. En este problema se utilizaron datos de disimilaridad entre 18 especies de flores a los que se les aplicó el algoritmo SMACOF y se determinó el número adecuado de dimensiones que es de 5 (Figura 2.1), se realizó su diagrama de Shepard para ese número de dimensiones, pero probando diferentes ajustes encontrando que el mejor es el de tipo ordinal (Figura 2.2) ya que su Stress-1 de 0.041 es de bueno a excelente.

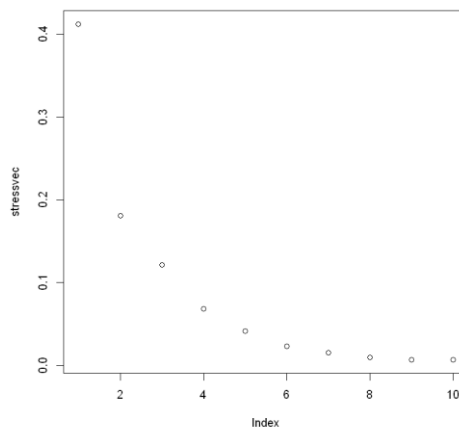


Figura 2.1: Stress-1 vs dimensionalidad, punto de cambio en las pendientes en 5.

Por último en la configuración de la solución en las primeras dos dimensiones se pudo observar que existen 4 grupos de flores que comparten similitudes (Figura 2.3), por ejemplo haciendo una rápida evaluación en Google pude apreciar que estas similitudes son por color o por el tipo de pétalo, por ejemplo, las del grupo de la derecha son rojas (Camelia, Fuchsia, Begonia), las del grupo de grande arriba son como de color rosa o lila y las de la izquierda son entre amarillos o de color blanco.

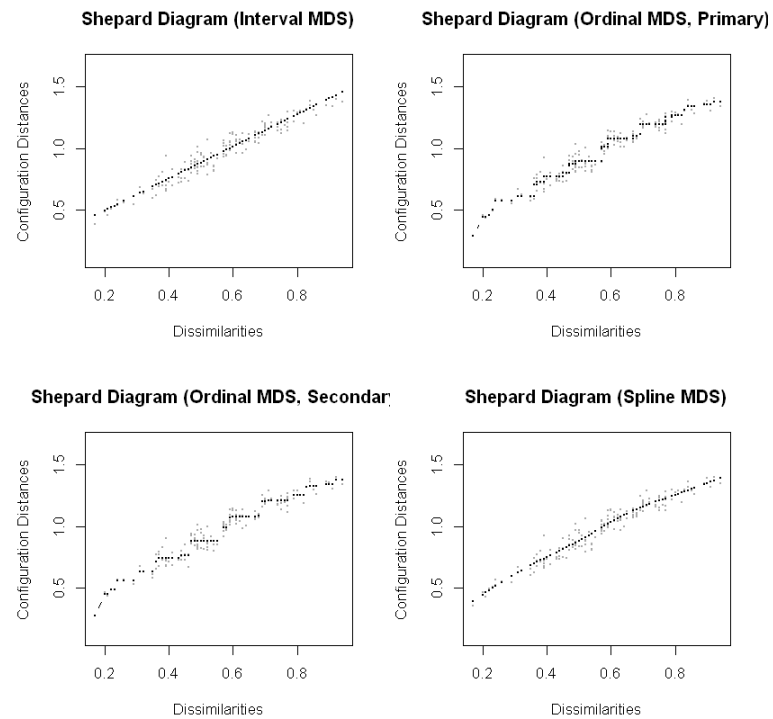


Figura 2.2: Diagramas de Shepard para 5 dimensiones con diferentes ajustes y sus respectivos Stress-1 son: 0.058 (Interval), 0.041 (ordinal1), 0.045 (ordinal2) y 0.056 (spline).

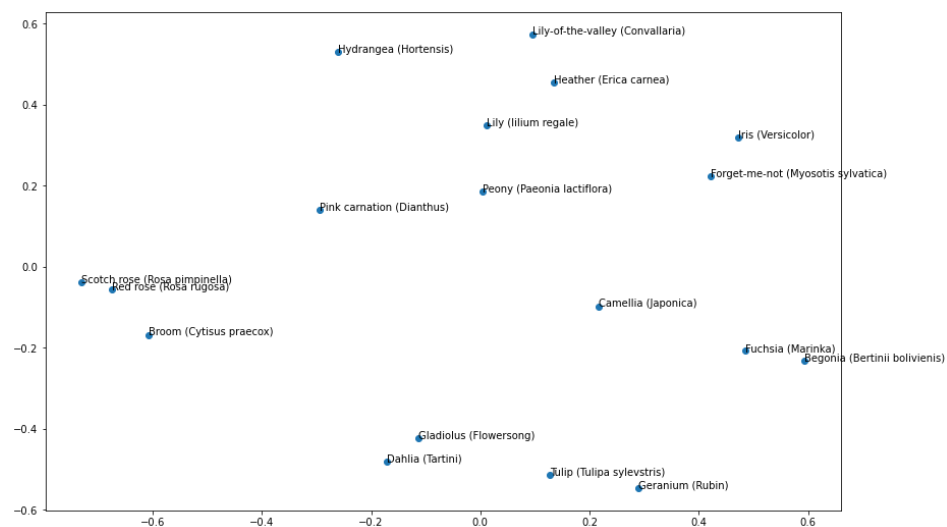


Figura 2.3: Primeras dos dimensiones de la solución.

3. Para este problema se utilizaron datos de preferencias de un estudio a 42 estudiantes (sujetos) ordenaron 15 tipos de desayuno (objetos) por orden de preferencia. Los desayunos son:

<i>Desayunos</i>
<i>Tostada sola (TS)</i>
<i>Tostada con mantequilla (TM)</i>
<i>Magdalena y margarina (MM)</i>
<i>Donas con jalea (DJ)</i>
<i>Tostada de canela (TC)</i>
<i>Bollos de arándanos y margarina (BAM)</i>
<i>Rosquillas y mantequilla (RM)</i>
<i>Tostada y mermelada (TMD)</i>
<i>Tostada con mantequilla y jalea (TMJ)</i>
<i>Tostada y margarina (TMG)</i>
<i>Pastel de canela (PC)</i>
<i>Pastas danesas (PD)</i>
<i>Donas glaseadas (DG)</i>
<i>Tarta de café (TC)</i>
<i>Bollo de maíz y mantequilla (BM)</i>



Para identificar como son discriminados los distintos desayunos se utilizó un modelo de unfolding condicionado por filas (sujetos, es decir que las preferencias entre los sujetos son independientes) del cual su valor de Stress-1 fue de 0.33 (pobre) y se obtuvo la solución en 2 dimensiones con una configuración circular para facilitar la interpretación.

Observe la Figura 3.1 y note que los desayunos (en color azul) están en la circunferencia y los estudiantes (color rojo) se encuentran dispersados dentro de la circunferencia, y mientras mas cerca se encuentren en algún punto de la circunferencia quiere decir que están mas a favor (por así decirlo) de ese desayuno (o desayunos) y mas alejados de los que no prefieren. Entonces se puede observar que del lado izquierdo parece haber una mayor concentración de sujetos principalmente cerca de PD y PC (Pastas danesas y pastel de canela) otros conjuntos que igual tienen cierta preferencia son los de PC, TCf y BAM del lado izquierdo y en el lado derecho los de TMG, PM y TM. Por otro lado, hay algunos que parece que casi nadie prefiere como TM, TS y TMD (Tostadas).

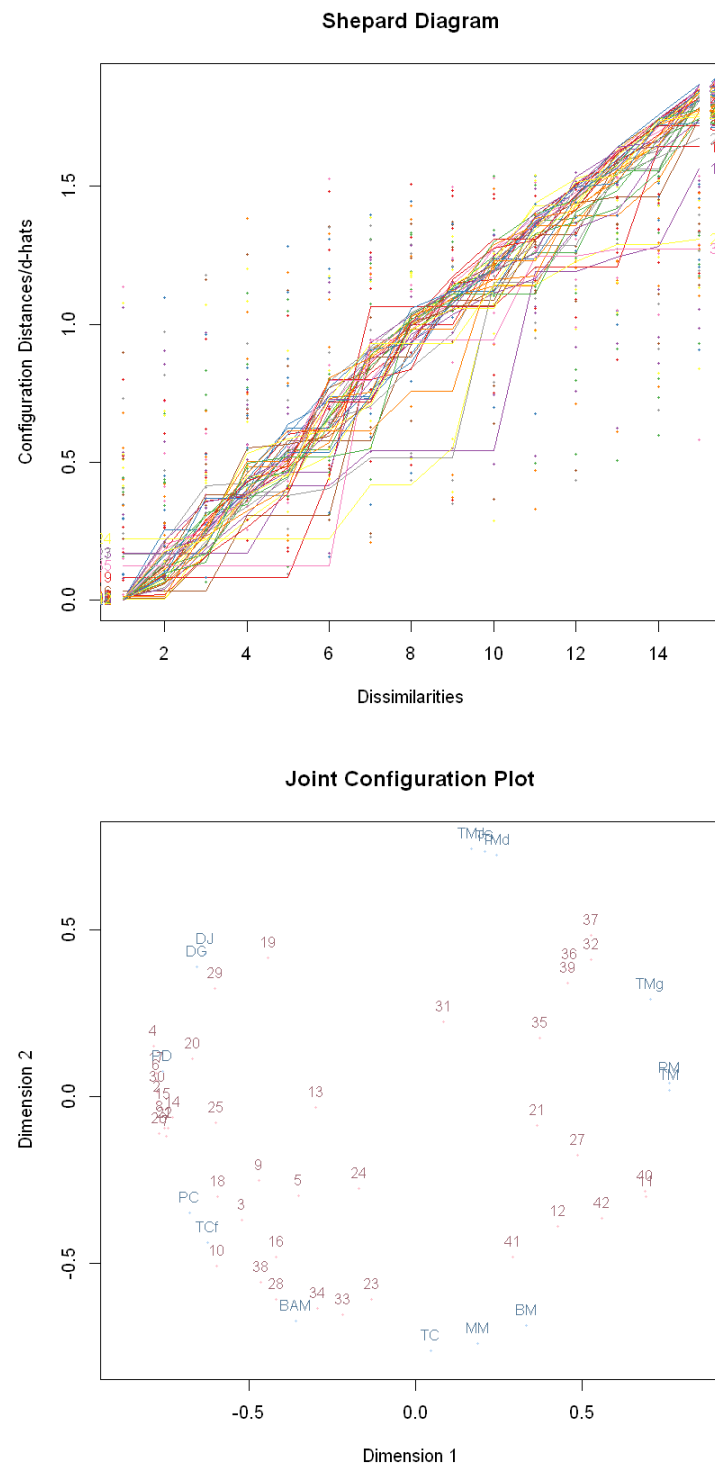


Figura 3.1: Parte superior diagrama de Shepard, parte inferior solución con en las primeras dos dimensiones con configuración circular.