Non-metric Multidimensional Scaling (NMDS)

- Fundamentally different than PCA, CA (and DCA)
- Ordination based on ranks
- Axes determined by rearranging points in an iterative way to find the “best” ordered solution

Major advantages of NMDS

- Data can be non-normal, or on arbitrary, discontinuous or questionable scales
- Avoids assumption of linear relationships among variables

The basic idea...

How would you represent these points in 2 dimensions?
One idea is do a projection...

Another is compress points to a surface

OBJECTIVE
close points stay close, far stay far

Which idea is better?
Both give correct results, but show different aspects of the data
The NMDS approach

- Uses a pre-specified number (small) of axes or dimensions (usually ≤ 3)
- Data is converted to a dissimilarity matrix among all pairs of samples
- Points in an initial ordination are rearranged in iterative fashion to find the lowest dimension solution
- A solution is found such that the rank order of distances between points in the ordination match the order of dissimilarity between the points

A Shepard diagram shows the relationship between dissimilarity and distance

Finding the best rank-order solution...

- Rank-order agreement of distances and dissimilarities is achieved by fitting a monotonic regression of the ordination distances onto the dissimilarities
- A monotonic regression is a non-parametric method of finding a least squares fit based on ranks of the dissimilarities
Shepard diagram with monotonic regression

Points that are off the monotonic regression line contribute to “stress” (or badness of fit)

Reducing “stress” in the solution

- Stress decreases as the rank-order agreement between distances and dissimilarities improves
- Thus, the aim is to find the solution with the lowest possible stress
- There is no algebraic method of finding the best solution; it must be found by an iterative search or trial-and-error optimization process
Basic steps for NMDS

1. Compute dissimilarities among the sample units
2. Specify the number of ordination dimensions (axes)
3. Generate an initial ordination of the sample units (starting configuration) with this number of axes
4. Calculate the distances between each pair of sample units in the current ordination
5. Perform a monotonic regression of the distances vs. dissimilarities (Shepherd plot)

Basic steps for NMDS, cont’d

6. Calculate the stress
7. Slightly move each point in the ordination in a way that will decrease stress
8. Repeat steps 4 – 7 until stress either approaches zero or stops decreasing

A real example of NMDS

Densities (per sq km) of 7 large mammal species in 9 areas of Rwenzori National Park, Uganda

<table>
<thead>
<tr>
<th>Area</th>
<th>Elephant</th>
<th>Warthog</th>
<th>Hippo</th>
<th>Water-buck</th>
<th>Kob</th>
<th>Topi</th>
<th>Buffalo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area 1</td>
<td>1</td>
<td>1</td>
<td>15</td>
<td>7</td>
<td>9</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>Area 2</td>
<td>3</td>
<td>3</td>
<td>28</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Area 3</td>
<td>3</td>
<td>8</td>
<td>1</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>Area 4</td>
<td>1</td>
<td>1</td>
<td>13</td>
<td>1</td>
<td>6</td>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td>Area 5</td>
<td>4</td>
<td>3</td>
<td>21</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>Area 6</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>44</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>Area 7</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>10</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>Area 8</td>
<td>0</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>78</td>
<td>30</td>
<td>22</td>
</tr>
<tr>
<td>Area 9</td>
<td>0</td>
<td>2</td>
<td>6</td>
<td>0</td>
<td>71</td>
<td>83</td>
<td>17</td>
</tr>
</tbody>
</table>
First create a dissimilarity matrix among the nine areas

- Use Bray-Curtis (Sørenson) measure
- Only need upper triangle, without the zero diagonals

<table>
<thead>
<tr>
<th></th>
<th>Area 1</th>
<th>Area 2</th>
<th>Area 3</th>
<th>Area 4</th>
<th>Area 5</th>
<th>Area 6</th>
<th>Area 7</th>
<th>Area 8</th>
<th>Area 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area 1</td>
<td>0.0000</td>
<td>0.3626</td>
<td>0.5121</td>
<td>0.2273</td>
<td>0.3053</td>
<td>0.5185</td>
<td>0.3086</td>
<td>0.7348</td>
<td>0.7500</td>
</tr>
<tr>
<td>Area 2</td>
<td>0.3626</td>
<td>0.0000</td>
<td>0.4384</td>
<td>0.4382</td>
<td>0.2392</td>
<td>0.7292</td>
<td>0.7248</td>
<td>0.5854</td>
<td>0.9462</td>
</tr>
<tr>
<td>Area 3</td>
<td>0.5121</td>
<td>0.4384</td>
<td>0.0000</td>
<td>0.4444</td>
<td>0.4021</td>
<td>0.5046</td>
<td>0.4158</td>
<td>0.7049</td>
<td>0.8130</td>
</tr>
<tr>
<td>Area 4</td>
<td>0.2273</td>
<td>0.4382</td>
<td>0.4444</td>
<td>0.0000</td>
<td>0.2988</td>
<td>0.5550</td>
<td>0.2152</td>
<td>0.6648</td>
<td>0.7297</td>
</tr>
<tr>
<td>Area 5</td>
<td>0.3053</td>
<td>0.2392</td>
<td>0.4021</td>
<td>0.2988</td>
<td>0.0000</td>
<td>0.4637</td>
<td>0.3954</td>
<td>0.7127</td>
<td>0.7817</td>
</tr>
<tr>
<td>Area 6</td>
<td>0.5185</td>
<td>0.7248</td>
<td>0.5046</td>
<td>0.5550</td>
<td>0.4637</td>
<td>0.0000</td>
<td>0.4343</td>
<td>0.4070</td>
<td>0.5124</td>
</tr>
<tr>
<td>Area 7</td>
<td>0.3086</td>
<td>0.4021</td>
<td>0.4021</td>
<td>0.2988</td>
<td>0.3954</td>
<td>0.4343</td>
<td>0.0000</td>
<td>0.6355</td>
<td>0.7023</td>
</tr>
<tr>
<td>Area 8</td>
<td>0.7348</td>
<td>0.4626</td>
<td>0.7094</td>
<td>0.5544</td>
<td>0.7274</td>
<td>0.4070</td>
<td>0.6195</td>
<td>0.0000</td>
<td>0.2254</td>
</tr>
<tr>
<td>Area 9</td>
<td>0.7500</td>
<td>0.4489</td>
<td>0.8230</td>
<td>0.7297</td>
<td>0.7817</td>
<td>0.5124</td>
<td>0.7023</td>
<td>0.2254</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Initial ordination generated randomly

Initial Shepard Diagram
Stress

- Stress of initial ordination is 4.183
  - This is high, reflecting the poor rank-order agreement of distances with dissimilarities at this stage.
- Each point is moved slightly
- Stress is recalculated and the process iterated until minimum stress is achieved

Shepard Diagram – Iteration 5

Shepard Diagram – Iteration 10
Shepard Diagram – Iteration 15

Shepard Diagram – Iteration 20

Shepard Diagram – Iteration 30
How the ordination evolved

How stress changed

Final Ordination

Once stress starts to level out, most SUs don’t change much in their position
There’s no way to know in advance how many dimensions you need

- To determine, NMDS must be run in a range of dimensionalities
- Most community data sets can be satisfactorily summarized with 2 or 3 NMDS axes; rarely are more needed

A Scree plot can guide determination of the number of dimensions

A break in slope, beyond which further reductions in stress are small, suggests dimensionality

No further reduction in stress

Stress guidelines for NMDS interpretability

Stresses up to 20 can be ecologically interpretable and useful

<table>
<thead>
<tr>
<th>Stress</th>
<th>Interpretability of NMDS</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 5</td>
<td>Excellent: no prospect of misinterpretation (rarely achieved)</td>
</tr>
<tr>
<td>5 - 10</td>
<td>Good: little danger of drawing false inferences</td>
</tr>
<tr>
<td>10 - 20</td>
<td>Fair: usable, but some distances will be misleading</td>
</tr>
<tr>
<td>&gt; 20</td>
<td>Poor: ordination may be dangerous to interpret</td>
</tr>
</tbody>
</table>