Statistics for the Sciences

Multidimensional Scaling

Xuemao Zhang
East Stroudsburg University

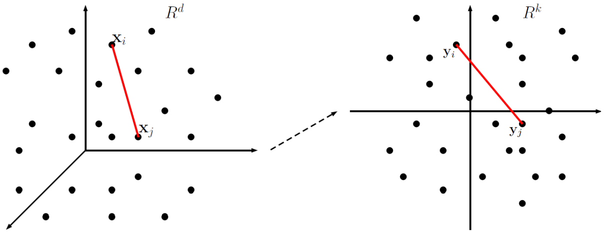
January 18, 2025

Outline

- Purpose of multidimensional scaling
- Multidimensional Scaling (MDS)
  - metric MDS
  - non-metric MDS
- Example
- Lab

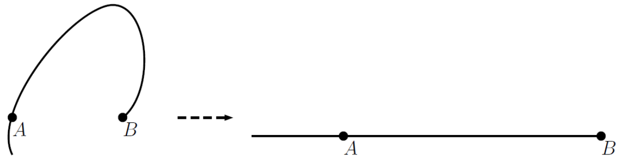
Purpose of multidimensional scaling

- Multidimensional scaling (MDS) is another technique on dimensionality reduction.
- Unlike PCA, MDS maps the original high dimensional space to a lower dimensional space, but does so in an attempt to **preserve pairwise distances**.
  - The space is usually two-dimensional, sometimes may be three-dimensional.



Purpose of multidimensional scaling

- The purpose of MDS is often just visualizing the data so it becomes easier for the user to explore and to understand their **structure**.
- The MDS algorithm starts with a distance matrix and uses it to reduce the data to two dimensions while preserving the pairwise distances.
  - PCA starts with the correlation matrix of data.



Multidimensional Scaling

- MDS methods include
  - Classical MDS (Metric MDS) or Principal Coordinates Analysis (PCoA),
  - Non-metric MDS which focuses on preserving the rank order of the dissimilarities rather than the actual distances.

Multidimensional Scaling

- Given p-dimensional data, how do we construct a **distance** matrix?

$$\boldsymbol{X} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}$$

Multidimensional Scaling

- **Distance**, **dissimilarity** and **similarity** (or **proximity**) are defined for any pair of objects in any space.
  - Each objects or subject is just a row (record) of observations.
- Let x, y, z be objects. In mathematics, a distance function (that gives a distance between two objects) is also called **metric**, satisfying
  - $d(x, y) \geq 0$
  - $d(x, y) = 0$ if and only if $x = y$
  - $d(x, y) = d(y, x)$
  - $d(x, z) \leq d(x, y) + d(y, z)$
- Given a set of dissimilarities, one can ask whether these values are distances.

Multidimensional Scaling

- Dissimilarity Matrix: $n \times n$ with the ij-th element d_{ij} measuring the distance (similarity or proximity) between the ith and the jth objects (or observations). D is symmetric.
  - Here p is the dimension of the data or the number of variables.
- Dissimilarity between points i and $i^{'}$: $\mathbf{x}_i$ and $\mathbf{x}_{i'}$ $(i, i^{'} = 1, \ldots, n)$:

$$d(\mathbf{x}_i, \mathbf{x}_{i'}) = \sum_{j=1}^{p} d_j(x_{ij}, x_{i'j}).$$

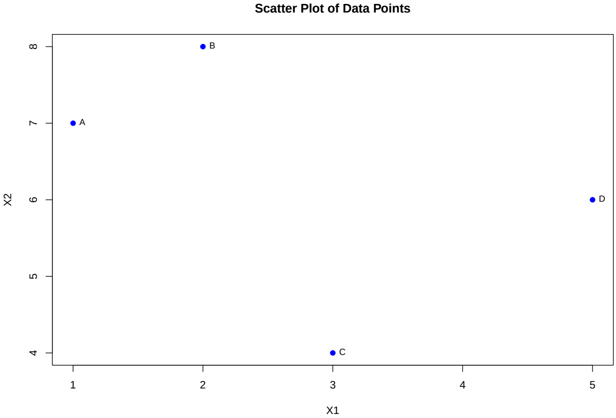
Multidimensional Scaling

- An example

```
##    x1 x2
## A  1  7
## B  2  8
## C  3  4
## D  5  6
```

Multidimensional Scaling

- A scatter plot of these data



Multidimensional Scaling

- Construct a dissimilarity matrix

```
## [1] "Distance matrix of the data:"

##          A        B        C        D
## A 0.000000 1.414214 3.605551 4.123106
## B 1.414214 0.000000 4.123106 3.605551
## C 3.605551 4.123106 0.000000 2.828427
## D 4.123106 3.605551 2.828427 0.000000
```

Multidimensional Scaling

- There are several types of distances. Consider two points $A(x_1, y_1)$ and $B(x_2, y_2)$
  - Euclidean distance $d(A, B) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$
  - Maximum Distance $d(A, B) = max(|x_1 - x_2|, |y_1 - y_2|)$
  - Manhattan Distance $d(A, B) = |x_1 - x_2| + |y_1 - y_2|$
  - Bray Distance $d(A, B) = \dfrac{|x_1 - x_2|}{|x_1| + |x_2|} + \dfrac{|y_1 - y_2|}{|y_1| + |y_2|}$
  - Minkowski Distance $d(A, B) = \left[(x_1 - x_2)^p + (y_1 - y_2)^p\right]^{1/p}, p > 0$
  - Binary Distance (Dissimilarity for binary data) $d(A, B) = (x_1 \text{ XOR } x_2) + (y_1 \text{ XOR } y_2)$

Multidimensional Scaling

The MDS problem:

- Assume a collection of p-dimensional n objects with Dissimilarity Matrix $\mathbf{L} = \{l_{ij}\}, i, j = 1, 2, \ldots, n$.
- Our objective is to represent them as points in 2-d Euclidean space, $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n \in \mathbb{R}^2$, such that

$$d_{ij} = d_{\boldsymbol{y}_i, \boldsymbol{y}_j} = l_{ij} \text{ or as close as possible for any } i, j.$$

- MDS
  - Called **metric MDS** if dissimilarities l_{ij} are quantitative.
  - Called **non-metric MDS** if dissimilarities l_{ij} are qualitative (e.g. ordinal).

Multidimensional Scaling

- The mathematical tool for MDS often involves eigen decomposition of the Dissimilarity matrix to transform the distance information into a lower-dimensional space while preserving the distances between data points.
- Solution to MDS is not unique.
  - If $\boldsymbol{Y}$ is a solution, then $\boldsymbol{Y}^* = \boldsymbol{Y} + \mathbf{c}, \mathbf{c} \in \mathbb{R}^2$ is also a solution.
  - $\mathbf{c}$ is just a location transformation.

Non-metric MDS

- In many applications of MDS, dissimilarities are known only by their **rank order**, and the spacing between successively ranked dissimilarities is of no interest or is unavailable.
  - Non-metric MDS focuses on preserving the rank order of the dissimilarities.
- Non-metric MDS: non-metric MDS seeks to find an optimal configuration $\boldsymbol{Y} \subset \mathbb{R}^2$ such that

$$f(l_{ij}) \approx d_{\boldsymbol{y}_i,\boldsymbol{y}_j} = d_{ij} \text{ for any } i, j.$$

  where f is a general monotonic function implicitly defined by the rank order of the original dissimilarities l_{ij}.
  - $f(l_{ij})$ generates **disparities**, which only preserve the order of the original dissimilarities l_{ij}. That is

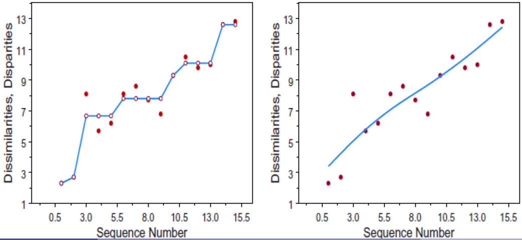
$$l_{ij} < l_{i'j'} \Leftrightarrow f(l_{ij}) < f(l_{i'j'})$$

Non-metric MDS

- Kruskal's non-metric MDS minimizes the stress-1

$$\text{stress-1}(f(I_{ij}), d_{ij}) = \left(\sum_{i<j} \frac{(f(I_{ij}) - d_{ij})^2}{\sum d_{ij}^2} \right)^{1/2}$$

- It quantifies the difference between the disparities $f(I_{ij})$ and the distances d_{ij} in $\mathbb{R}^2$.
- The function f works as if it were a regression curve
  - approximated dissimilarities d_{ij} treated as response,
  - disparities $f(I_{ij})$ serves as estimate average of the response, and
  - the order of dissimilarities as explanatory
  - In the following figure, Left panel: Isotonic regression. right panel: Monotone spline. For the red points, the vertical axis is the dissimilarity I_{ij}, whereas for the fitted blue points, the vertical axis is the disparity $f(I_{ij})$.



Non-metric MDS

- Process of Non-metric MDS
  - Initialization: Start with an initial configuration of points in the lower-dimensional space like $\mathbb{R}^2$.
  - Compute Distances: Calculate the Euclidean distances d_{ij} between points in the current configuration.
  - Monotonic Regression: Find the disparities $f(l_{ij})$ that best match the distances d_{ij} while preserving the rank order of l_{ij}.
  - Optimization: Adjust the configuration of points (the coordinates of $\boldsymbol{y}$) to minimize the stress function in the last slide. The goal is to find the configuration where the distances d_{ij} are as close as possible to the disparities $f(l_{ij})$.

Example 1: Metric MDS

- `lemminvert.csv`: Lemmens et al. (2015) did a detailed study of various biotic communities in artificial ponds in Belgium. They sampled 28 ponds that represented different types of management, a combination of fish farming strategies (no fish, farming young fish, low intensity management, no management), and drainage frequencies ($>$ 10 years ago, occasional, annual). They also quantified taxon abundances for fish, zooplankton, and macro-invertebrates (different families and species within some groups) and covers of submerged, floating, and emergent vegetation. The macroinvertebrate dataset only included 23 ponds.
- See analysis with R here: https://mjkeough.github.io/examples/lemmpcoa.nb.html

```
##  [1] "1J1" "1J2" "1J3" "1J5" "1J6" "1J7" "BK1" "BK2" "BK5" "BK6"
## [13] "O2"  "O3"  "O4"  "O5"  "O6"  "O7"  "V1"  "V2"  "V3"  "V4"
```

Example 1: Metric MDS

- The first 3 variables are pond classification variables

```
## 'data.frame':    23 obs. of  32 variables:
##  $ site     : Factor w/ 23 levels "1J1","1J2","1J3",..: 1 2 3 4 5 6 7 8 9 10 ...
##  $ manag    : Factor w/ 4 levels "li","nf","nm",..: 4 4 4 4 4 4 2 2 2 2 ...
##  $ managsymb: int  0 0 0 0 0 0 1 1 1 1 ...
##  $ caen     : int  0 6 0 5 0 3 31 36 0 16 ...
##  $ baet     : int  158 13 20 98 3 605 1571 494 73 90 ...
##  $ acrol    : int  0 0 0 0 0 0 0 0 1 0 ...
##  $ lymn     : int  3 0 3 10 9 7 763 109 46 15 ...
##  $ plan     : int  5 50 16 206 14 19 1164 1035 64 421 ...
##  $ sphaer   : int  0 0 0 2 0 0 80 147 0 0 ...
##  $ vivip    : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ corix    : int  671 95 8 98 23 9 730 389 239 25 ...
##  $ nauc     : int  0 0 0 0 0 0 0 2 2 0 ...
##  $ nepid    : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ noto     : int  0 0 0 0 0 0 3 1 0 0 ...
##  $ pleid    : int  0 0 0 0 0 0 20 141 5 0 ...
##  $ gamm     : int  0 0 1 2 1 0 4 0 4 1 ...
##  $ asell    : int  0 0 0 4 0 0 0 0 0 1 ...
##  $ cerat    : int  4 7 9 5 0 7 88 3 6 20 ...
##  $ chaob    : int  23 37 60 20 3 5 68 0 204 22 ...
##  $ chiro    : int  222 322 240 26 60 394 1236 62 569 69 ...
##  $ culic    : int  19 17 12 4 3 19 12 0 24 21 ...
##  $ cylind   : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ dixid    : int  40 52 30 2 5 0 4 6 47 38 ...
##  $ empid    : int  0 0 0 0 1 0 0 0 0 5 ...
##  $ ephyd    : int  0 0 0 0 1 0 0 0 0 0 ...
##  $ limon    : int  8 1 25 9 2 0 8 0 0 10 ...
##  $ psych    : int  0 0 0 0 1 0 0 0 0 0
```

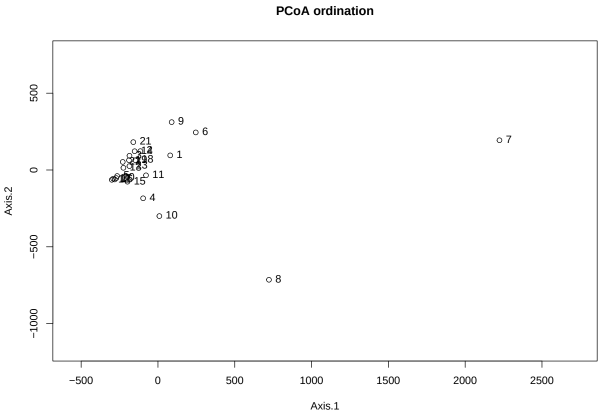
Example 1: Metric MDS

- PCoA Analysis

```
## $correction
## [1] "none" "1"
##
## $note
## [1] "There were no negative eigenvalues. No correction was applied"
##
## $values
##     Eigenvalues Relative_eig Broken_stick Cumul_eig Cumul_br_stick
## 1  6.284770e+06 7.820467e-01  0.167764239 0.7820467      0.1677642
## 2  9.438104e+05 1.174433e-01  0.122309693 0.8994900      0.2900739
## 3  4.249575e+05 5.287968e-02  0.099582420 0.9523696      0.3896564
## 4  2.622459e+05 3.263263e-02  0.084430905 0.9850023      0.4740873
## 5  7.088309e+04 8.820352e-03  0.073067269 0.9938226      0.5471545
## 6  2.045955e+04 2.545889e-03  0.063976360 0.9963685      0.6111309
## 7  1.124622e+04 1.399426e-03  0.056400602 0.9977679      0.6675315
## 8  5.798579e+03 7.215475e-04  0.049907096 0.9984895      0.7174386
## 9  5.194204e+03 6.463420e-04  0.044225278 0.9991358      0.7616639
## 10 2.820873e+03 3.510159e-04  0.039174773 0.9994868      0.8008386
## 11 1.844753e+03 2.295523e-04  0.034629318 0.9997164      0.8354680
## 12 1.287003e+03 1.601485e-04  0.030497087 0.9998765      0.8659650
## 13 5.023922e+02 6.251528e-05  0.026709208 0.9999391      0.8926742
## 14 2.379086e+02 2.960421e-05  0.023212704 0.9999687      0.9158870
## 15 1.646356e+02 2.048647e-05  0.019965951 0.9999891      0.9358529
## 16 4.038504e+01 5.025321e-06  0.016935648 0.9999942      0.9527886
## 17 2.848683e+01 3.544765e-06  0.014094739 0.9999977      0.9668833
## 18 8.453043e+00 1.051856e-06  0.011420942 0.9999988      0.9783042
## 19 4.215181e+00 5.245170e-07  0.008895690 0.9999993      0.9871999
## 20 3.346910e+00 4.164734e-07  0.006503345 0.9999997      0.9937033
## 21 1.664062e+00 2.070680e-07  0.004230618 0.9999999      0.9979339
## 22 6.604191e-01 8.217940e-08  0.002066116 1.0000000      1.0000000
##
## $vectors
##        Axis.1     Axis.2    Axis.3      Axis.4      Axis.5       Axis.6
## 1   79.740869   95.10472 498.99343  215.307907  -38.041790 -21.18016513
## 2 -114.437233  122.79086  36.06851 -115.990258  -23.898179 -39.23757341
## 3 -185.375472   92.69808 -47.57754  -84.558400  -12.730992   2.97584395
## 4  -96.499521 -184.16382 -11.40450   17.971440    8.590388  -0.03897887
```

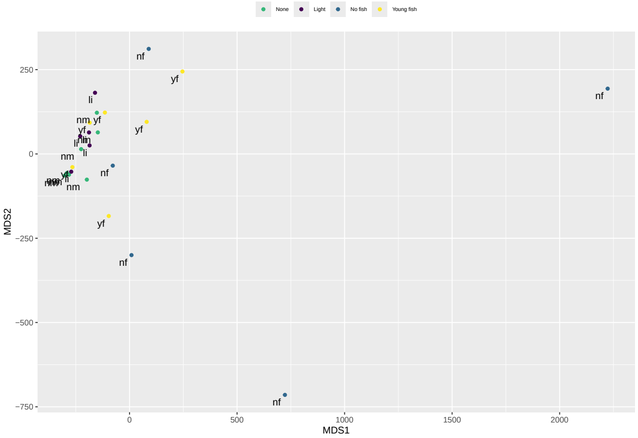
Example 1: Metric MDS

- biplot



Example 1: Metric MDS

- biplot



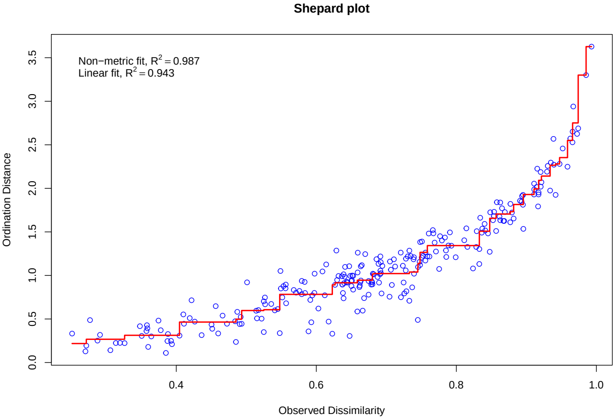
Example 2: Non-Metric MDS

- NDMS using `bray` distance https://mjkeough.github.io/examples/lemmmds.nb.html

```
## Run 0 stress 0.1156702
## Run 1 stress 0.1153229
## ... New best solution
## ... Procrustes: rmse 0.01893663  max resid 0.07573146
## Run 2 stress 0.115323
## ... Procrustes: rmse 0.0002871722  max resid 0.001111882
## ... Similar to previous best
## Run 3 stress 0.1153229
## ... New best solution
## ... Procrustes: rmse 0.0001334567  max resid 0.0005117371
## ... Similar to previous best
## Run 4 stress 0.115323
## ... Procrustes: rmse 0.0002058442  max resid 0.0008060438
## ... Similar to previous best
## Run 5 stress 0.1153229
## ... Procrustes: rmse 0.0001425949  max resid 0.0005563484
## ... Similar to previous best
## Run 6 stress 0.1172332
## Run 7 stress 0.1156703
## ... Procrustes: rmse 0.01890827  max resid 0.07510586
## Run 8 stress 0.11567
## ... Procrustes: rmse 0.01893865  max resid 0.07545109
## Run 9 stress 0.1156703
## ... Procrustes: rmse 0.0189112  max resid 0.07523154
## Run 10 stress 0.1172331
## Run 11 stress 0.1156701
```

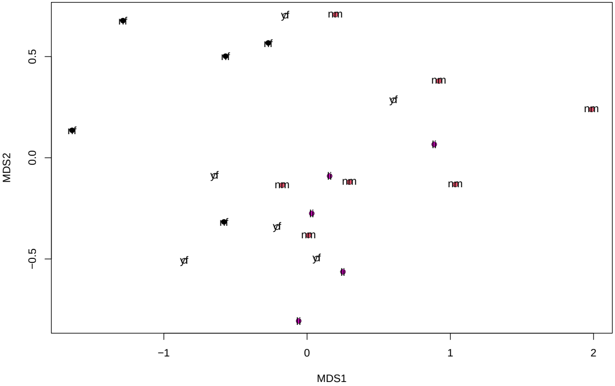
Example 2: Non-Metric MDS

- Stress plot



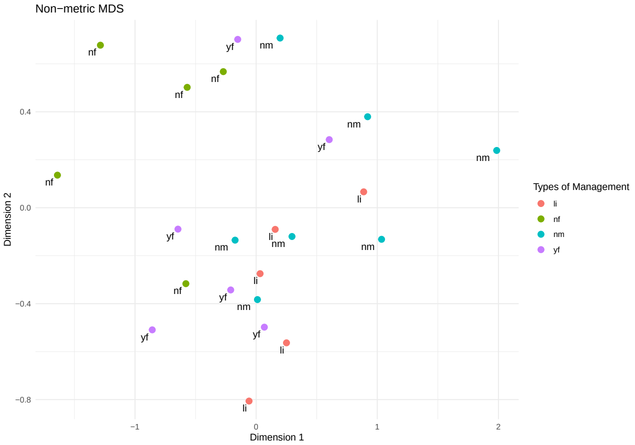
Example 2: Non-Metric MDS

- biplot



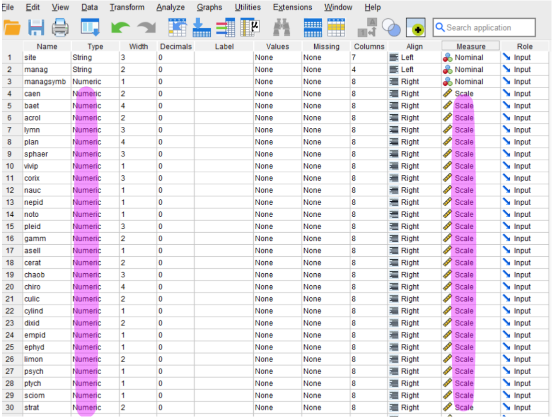
Example 2: Non-Metric MDS

- biplot



Lab

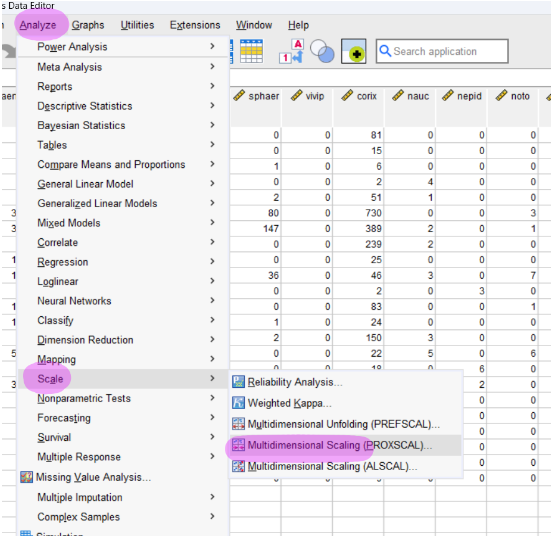
- After importing data, change the measures of all numercal variables to `Scale`



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	site	String	3	0		None	None	7	Left	Nominal	Input
2	manag	String	2	0		None	None	4	Left	Nominal	Input
3	managsymb	Numeric	1	0		None	None	8	Right	Nominal	Input
4	caen	Numeric	2	0		None	None	8	Right	Scale	Input
5	baet	Numeric	4	0		None	None	8	Right	Scale	Input
6	acrol	Numeric	2	0		None	None	8	Right	Scale	Input
7	lymn	Numeric	3	0		None	None	8	Right	Scale	Input
8	plan	Numeric	4	0		None	None	8	Right	Scale	Input
9	sphaer	Numeric	3	0		None	None	8	Right	Scale	Input
10	vivip	Numeric	1	0		None	None	8	Right	Scale	Input
11	corix	Numeric	3	0		None	None	8	Right	Scale	Input
12	nauc	Numeric	1	0		None	None	8	Right	Scale	Input
13	nepid	Numeric	1	0		None	None	8	Right	Scale	Input
14	noto	Numeric	1	0		None	None	8	Right	Scale	Input
15	pleid	Numeric	3	0		None	None	8	Right	Scale	Input
16	gamm	Numeric	2	0		None	None	8	Right	Scale	Input
17	asell	Numeric	1	0		None	None	8	Right	Scale	Input
18	cerat	Numeric	2	0		None	None	8	Right	Scale	Input
19	chaob	Numeric	3	0		None	None	8	Right	Scale	Input
20	chiro	Numeric	4	0		None	None	8	Right	Scale	Input
21	culic	Numeric	2	0		None	None	8	Right	Scale	Input
22	cylind	Numeric	1	0		None	None	8	Right	Scale	Input
23	dixid	Numeric	2	0		None	None	8	Right	Scale	Input
24	empid	Numeric	1	0		None	None	8	Right	Scale	Input
25	ephyd	Numeric	1	0		None	None	8	Right	Scale	Input
26	limon	Numeric	2	0		None	None	8	Right	Scale	Input
27	psych	Numeric	1	0		None	None	8	Right	Scale	Input
28	ptych	Numeric	1	0		None	None	8	Right	Scale	Input
29	sciom	Numeric	1	0		None	None	8	Right	Scale	Input
30	strat	Numeric	2	0		None	None	8	Right	Scale	Input

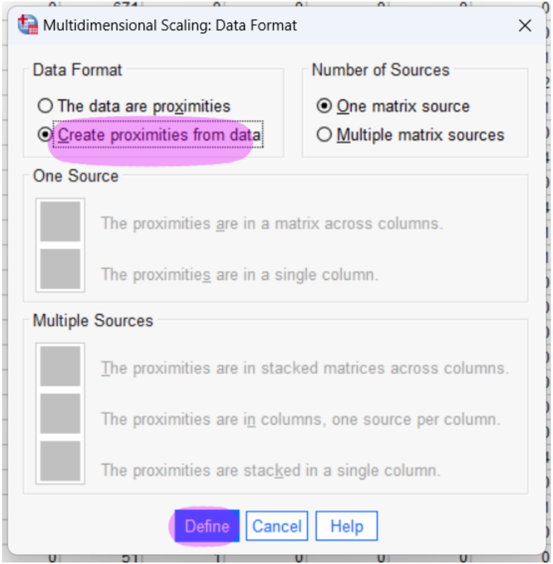
Lab

- Go to `Analyze` → `Scale` → `Multidimensional Scaling (ProxSCAL)...`
  - `PROXSCAL` is for metric and non-metric MDS



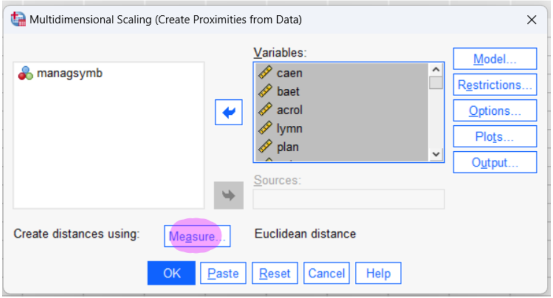
Lab

- We have raw data only, so we need to construct a distance matrix



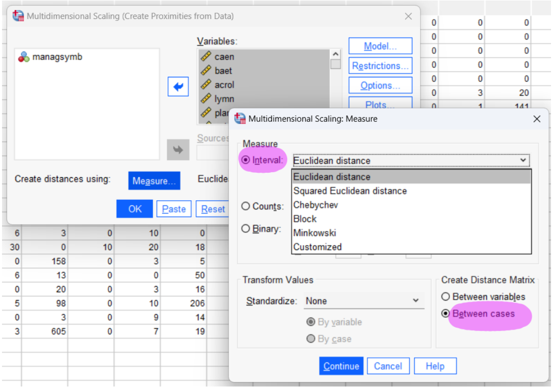
Lab

- Move all numerical variables to the `Variables` box



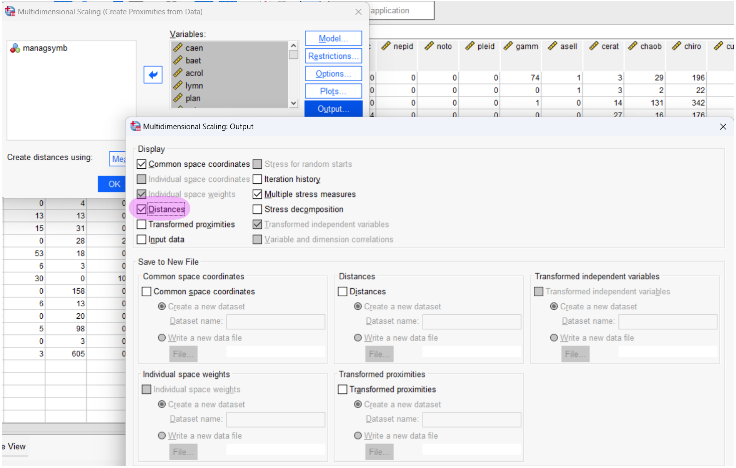
Lab

- Suppose we use Euclidean distance
  - Make sure we create distance matrix **Between cases**



Lab

- Output options



Lab

- Stress Value: This measures the goodness-of-fit of the MDS solution. Lower values indicate a better fit.
  - Typically, stress values below 0.1 indicate a good fit.

Stress and Fit Measures

Normalized Raw Stress	.00753
Stress-I	.08678[a]
Stress-II	.11662[a]
S-Stress	.01063[b]
Dispersion Accounted For (D.A.F.)	.99247
Tucker's Coefficient of Congruence	.99623

PROXSCAL minimizes Normalized Raw Stress.

a. Optimal scaling factor = 1.008.

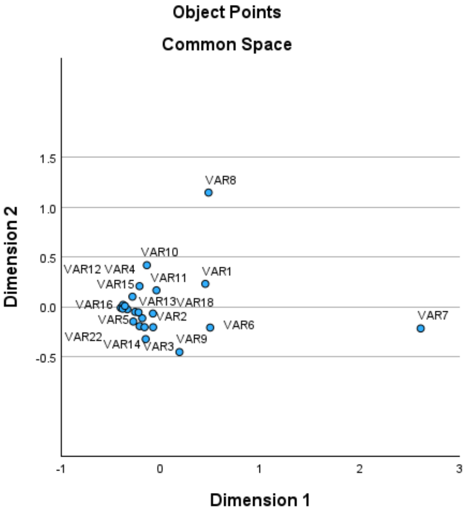
b. Optimal scaling factor = .994.

Lab

Final Coordinates

	Dimension	
	1	2
VAR1 Case 1	.445	.231
VAR2 Case 2	-.079	-.206
VAR3 Case 3	-.215	-.194
VAR4 Case 4	-.216	.208
VAR5 Case 5	-.338	-.025
VAR6 Case 6	.494	-.209
VAR7 Case 7	2.610	-.218
VAR8 Case 8	.478	1.148
VAR9 Case 9	.185	-.455
VAR10 Case 10	-.142	.418
VAR11 Case 11	-.046	.166
VAR12 Case 12	-.404	-.010
VAR13 Case 13	-.261	-.051
VAR14 Case 14	-.166	-.205
VAR15 Case 15	-.287	.102
VAR16 Case 16	-.382	.020
VAR17 Case 17	-.386	-.019
VAR18 Case 18	-.081	-.067
VAR19 Case 19	-.188	-.115
VAR20 Case 20	-.362	.008
VAR21 Case 21	-.154	-.325
VAR22 Case 22	-.279	-.148
VAR23 Case 23	-.226	-.055

Lab


Object Points
Common Space
Dimension 2
Dimension 1
1.5
1.0
0.5
0.0
-0.5
-1
0
1
2
3
VAR8
VAR10
VAR12 VAR4
VAR15
VAR11
VAR1
VAR16
VAR13VAR18
VAR2
VAR5
VAR6
VAR7
VAR22
VAR14
VAR3
VAR9

Lab

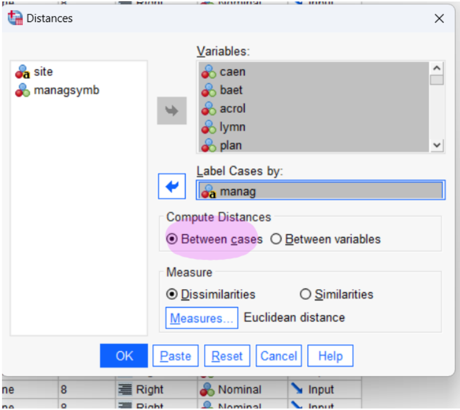
- Distance matrix

Distances

	VAR1 Case 1	VAR2 Case 2	VAR3 Case 3	VAR4 Case 4	VAR5 Case 5	VAR6 Case 6	VAR7 Case 7	VAR8 Case 8	VAR9 Case 9	VAR10 Case 10	VAR11 Case 11	VAR12 Case 12	VAR13 Case 13	VAR14 Case 14	VAR15 Case 15	VAR16 Case 16	VAR17 Case 17	VAR18 Case 18	VAR19 Case 19	VAR20 Case 20	VAR21 Case 21	VAR22 Case 22	VAR23 Case 23
VAR1 Case 1	.000																						
VAR2 Case 2	.682	.000																					
VAR3 Case 3	.786	.137	.000																				
VAR4 Case 4	.662	.435	.402	.000																			
VAR5 Case 5	.824	.316	.210	.263	.000																		
VAR6 Case 6	.442	.573	.710	.824	.853	.000																	
VAR7 Case 7	2.211	2.689	2.826	2.858	2.955	2.116	.000																
VAR8 Case 8	.918	1.464	1.511	1.169	1.429	1.358	2.532	.000															
VAR9 Case 9	.733	.363	.478	.775	.678	.395	2.436	1.630	.000														
VAR10 Case 10	.617	.627	.617	.223	.484	.894	2.825	.958	.932	.000													
VAR11 Case 11	.495	.373	.398	.176	.349	.657	2.683	1.113	.662	.270	.000												
VAR12 Case 12	.883	.380	.264	.288	.068	.920	3.022	1.456	.739	.502	.400	.000											
VAR13 Case 13	.760	.239	.151	.262	.082	.771	2.876	1.408	.602	.484	.305	.149	.000										
VAR14 Case 14	.751	.087	.051	.416	.249	.660	2.776	1.499	.431	.624	.390	.308	.181	.000									
VAR15 Case 15	.744	.371	.305	.127	.137	.841	2.915	1.296	.730	.348	.250	.162	.155	.330	.000								
VAR16 Case 16	.853	.377	.271	.250	.062	.905	3.001	1.418	.740	.465	.366	.038	.140	.312	.125	.000							
VAR17 Case 17	.868	.359	.245	.283	.048	.901	3.003	1.452	.719	.501	.387	.020	.129	.288	.156	.039	.000						
VAR18 Case 18	.605	.138	.185	.307	.261	.592	2.695	1.338	.470	.490	.236	.328	.180	.161	.267	.313	.309	.000					
VAR19 Case 19	.722	.142	.084	.324	.175	.689	2.800	1.428	.505	.535	.315	.240	.097	.093	.238	.236	.220	.117	.000				
VAR20 Case 20	.838	.355	.250	.247	.041	.884	2.981	1.416	.717	.465	.354	.046	.118	.290	.120	.023	.036	.291	.213	.000			
VAR21 Case 21	.817	.141	.144	.536	.352	.658	2.766	1.603	.363	.743	.502	.402	.294	.121	.447	.413	.384	.268	.213	.393	.000		
VAR22 Case 22	.817	.208	.078	.361	.137	.775	2.890	1.501	.556	.583	.391	.187	.099	.126	.250	.197	.168	.213	.096	.177	.216	.000	
VAR23 Case 23	.730	.210	.139	.263	.117	.736	2.841	1.394	.573	.481	.285	.184	.035	.161	.169	.173	.164	.145	.070	.151	.279	.107	.000

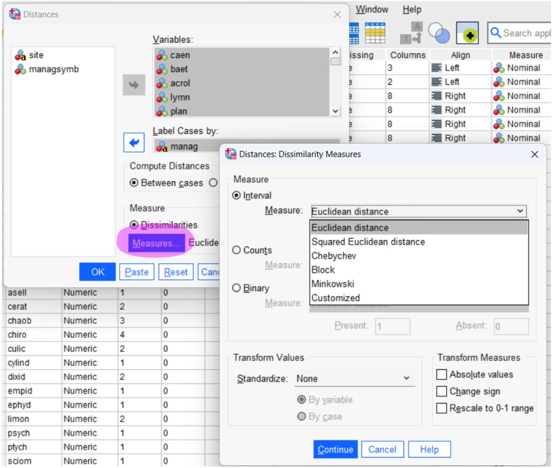
Lab

- Another method is that we start with the input of a **Distance Matrix**
  - For example, in SPSS, go to `Analyze` → `Correlate` → `Distances`.
  - In the dialog box, choose the variables you want to calculate distances for (we select all except the first three variables).
  - We label the rows by `manag`



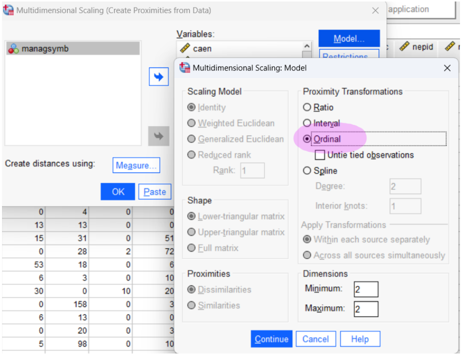
Lab

- Select the type of distance measure you want to use (e.g., Euclidean).
    - Click `OK` to compute the distance matrix. And then export the matrix to an Excel file → Convert to a CSV file and then import this distance matrix



Lab

- To conduct non-metric MDS with SPSS, we can still click `Analyze` $\rightarrow$ `Scale` $\rightarrow$ `Multidimensional Scaling (ProxSCAL)....`
- In the model,we choose `Ordinal` Proximity transformation



- It is better to have a distance matrix as input.SPSS does not support many distance functions, we will need to prepare our distance matrix using another software.
  - for example, function `vegdist` in R library `vegan`.

License