Ordination: Multidimensional Scaling

ST 558: Multivariate Analytics

Module 9

Lecture 1

Ordination

Ordination is a broad category of analysis techniques that aim to represent multivariate data in lower dimensions, generally with the goal of visualization.

We will consider ***multidimensional scaling*** as a particular ordination technique; PCA is another possible ordination technique, and there are many others.

Multidimensional Scaling

- *Multidimensional scaling* (MDS), sometimes also called *principal coordinate analysis*, is an ordination method, used to produce a low-dimensional representation of the observations that captures most of the information in the pairwise distance or dissimilarity matrix.
- The main purpose of performing MDS is to be able to visualize high-dimensional data.
- If only the ranks of the dissimilarities are known or used, the method is called non-metric multidimensional scaling (NMS or NMDS).
- If the actual distance values are used, the method is called metric multidimensional scaling.

Multidimensional Scaling

- Setting: Observations on p variables obtained for n observation units, giving data

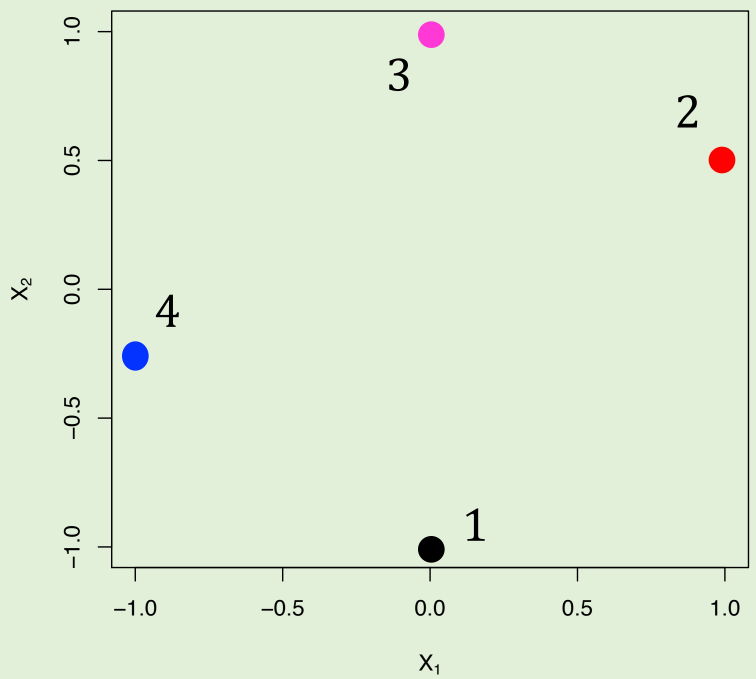
$$\mathbf{X}_i = \left[X_{i1}, X_{i2}, \ldots, X_{ip}\right]^T \text{for } i = 1, \ldots, n.$$

  - As with clustering, we truly only need a distance matrix $D = \left[d_{ij} = d\left(\mathbf{X}_i, \mathbf{X}_j\right)\right]$ that gives the pairwise distances/dissimiliarities between observations $\mathbf{X}_i$ and $\mathbf{X}_j$.

- Goal: Find a q-dimensional representation for the observations so that the distances $d_{i,j}^{(q)}$ between the q-dimensional representations $\mathbf{Z}_i$ and $\mathbf{Z}_j$ of the points $\mathbf{X}_i$ and $\mathbf{X}_j$ are as close as possible to the original distances, or at least have a similar ordering.

Multidimensional Scaling: Example

Example: Consider the following simple example, with n = 4 observations in p = 2 dimensions:



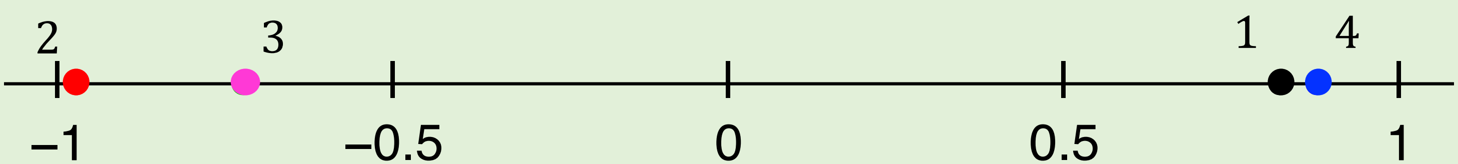
Multidimensional Scaling: Example

Example: The (Euclidean) distance matrix for these points is:

	1	**2**	**3**	**4**
1	0.00	1.80	2.00	1.25
2	1.80	0.00	1.12	2.14
3	2.00	1.12	0.00	1.60
4	1.25	2.14	1.60	0.00

Multidimensional Scaling: Example

<u>Example:</u> We can use multidimensional scaling to find a representation of these points in ***one*** dimension that best preserves the relative distances:



Multidimensional Scaling: Example

Example: The (Euclidean) distance matrix for the original points is:

	1	2	3	4
1	0.00	1.80	2.00	1.25
2	1.80	0.00	1.12	2.14
3	2.00	1.12	0.00	1.60
4	1.25	2.14	1.60	0.00

The (Euclidean) distance matrix for the points in this one-dimensional representation is:

	1	2	3	4
1	0.00	1.79	1.55	0.06
2	1.79	0.00	0.25	1.85
3	1.55	0.25	0.00	1.60
4	0.06	1.85	1.60	0.00

Multidimensional Scaling: Example

Example: The (Euclidean) distance matrix for the original points is:

	1	2	3	4
1	0.00	1.80	2.00	**1.25**
2	1.80	0.00	**1.12**	2.14
3	2.00	1.12	0.00	1.60
4	1.25	2.14	1.60	0.00

The (Euclidean) distance matrix for the points in dimensional representation is:

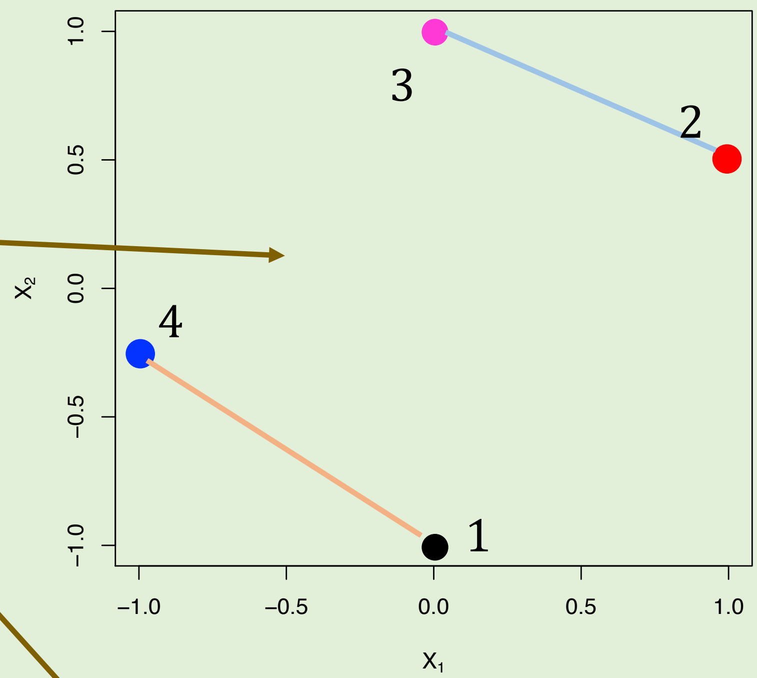
	1	2	3	4
1	0.00	1.79	1.55	**0.06**
2	1.79	0.00	**0.25**	1.85
3	1.55	0.25	0.00	1.60
4	0.06	1.85	1.60	0.00

The largest discrepancies between the original 2-dimensional distances and the new 1-dimensional distances are for these highlighted elements.

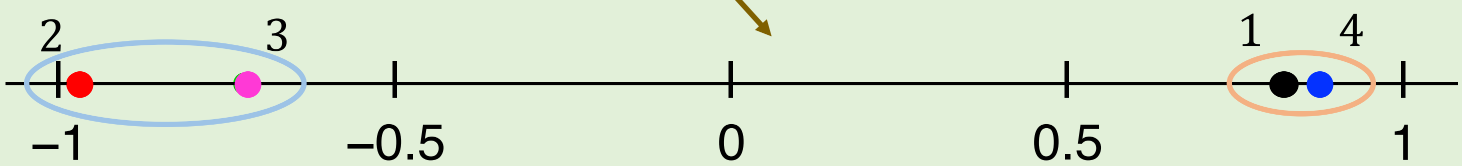
Multidimensional Scaling: Example

Example:

	1	2	3	4
1	0.00	1.80	2.00	**1.25**
2	1.80	0.00	**1.12**	2.14
3	2.00	1.12	0.00	1.60
4	1.25	2.14	1.60	0.00



	1	2	3	4
1	0.00	1.79	1.55	**0.06**
2	1.79	0.00	**0.25**	1.85
3	1.55	0.25	0.00	1.60
4	0.06	1.85	1.60	0.00



Multidimensional Scaling: Example

Example: We can also use MDS to reconstruct a map, given only distances between points. The flight distances between 11 US cities are given below:

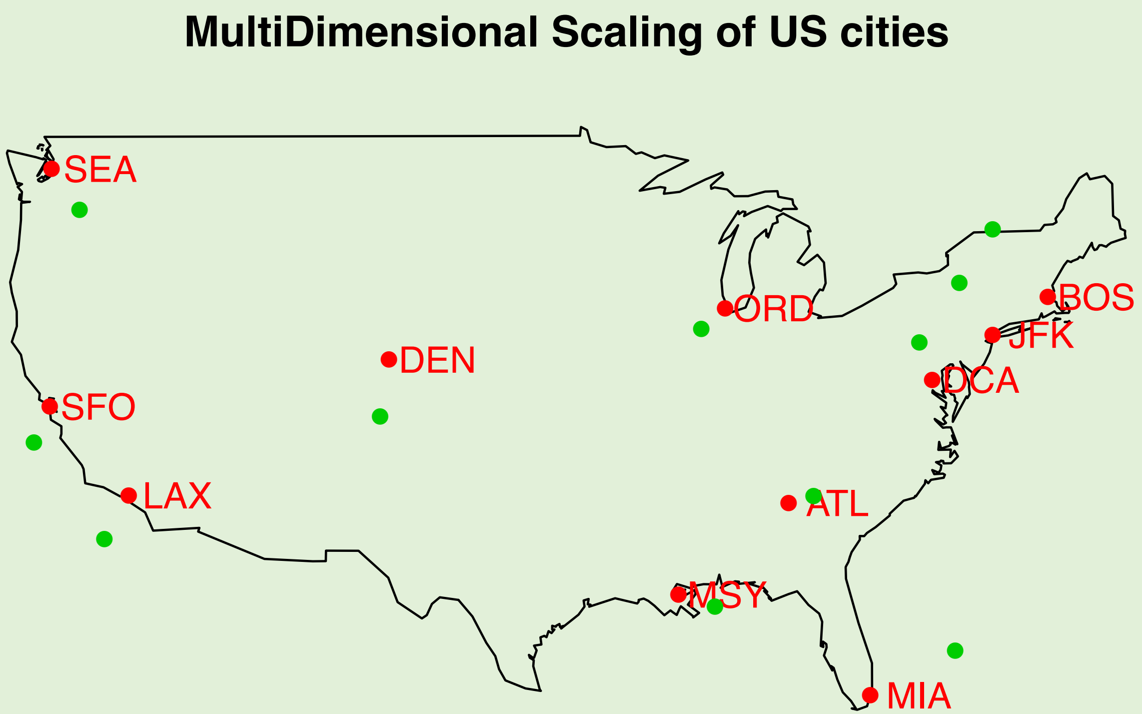
	ATL	BOS	ORD	DCA	DEN	LAX	MIA	JFK	SEA	SFO	MSY
ATL	0	934	585	542	1209	1942	605	751	2181	2139	424
BOS	934	0	853	392	1769	2601	1252	183	2492	2700	1356
ORD	585	853	0	598	918	1748	1187	720	1736	1857	830
DCA	542	392	598	0	1493	2305	922	209	2328	2442	964
DEN	1209	1769	918	1493	0	836	1723	1636	1023	951	1079
LAX	1942	2601	1748	2305	836	0	2345	2461	957	341	1679
MIA	605	1252	1187	922	1723	2345	0	1092	2733	2594	669
JFK	751	183	720	209	1636	2461	1092	0	2412	2577	1173
SEA	2181	2492	1736	2328	1023	957	2733	2412	0	681	2101
SFO	2139	2700	1857	2442	951	341	2594	2577	681	0	1925
MSY	424	1356	830	964	1079	1679	669	1173	2101	1925	0

Multidimensional Scaling: Example

Example: The true locations of these cities are shown below:

Multidimensional Scaling: Example

Example: The green dots are the locations assigned by multidimensional scaling:



Multidimensional Scaling: Classical Solution

The classical solution to the multidimensional scaling problem is also sometimes called <u>*principal coordinate analysis*</u>.

The solution is found as follows:

1. Let $a_{ij} = -\frac{1}{2}d_{ij}^2$
2. Define $b_{ij} = a_{ij} - \bar{a}_{i\cdot} - \bar{a}_{\cdot j} + \bar{a}_{\cdot\cdot}$, where

$$\bar{a}_{i\cdot} = \frac{1}{n}\sum_{j=1}^{n} a_{ij} \qquad \bar{a}_{\cdot j} = \frac{1}{n}\sum_{i=1}^{n} a_{ij} \qquad \bar{a}_{\cdot\cdot} = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}$$

Multidimensional Scaling: Classical Solution

(Classical MDS Solution, continued:)

3. Find the eigendecomposition of $\mathbf{B} = [b_{ij}]$, letting
   - $\lambda_1 > \lambda_2 > \ \ldots > \lambda_n$ be the eigenvalues, and
   - $\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_n$ be the corresponding $(n\times1)$ eigenvectors.
4. Define

$$\underset{(n\times q)}{\mathbf{Z}} = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \sqrt{\lambda_1}\mathbf{v}_1 & \sqrt{\lambda_2}\mathbf{v}_2 & \cdots & \sqrt{\lambda_q}\mathbf{v}_q \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} = \begin{bmatrix} \mathbf{Z}_1^T \\ \mathbf{Z}_2^T \\ \vdots \\ \mathbf{Z}_n^T \end{bmatrix}$$

Multidimensional Scaling: Classical Solution

Then the rows of

$$\underset{(n\times q)}{\mathbf{Z}} = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \sqrt{\lambda_1}\mathbf{v}_1 & \sqrt{\lambda_2}\mathbf{v}_2 & \cdots & \sqrt{\lambda_q}\mathbf{v}_q \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} = \begin{bmatrix} \mathbf{Z}_1^T \\ \mathbf{Z}_2^T \\ \vdots \\ \mathbf{Z}_n^T \end{bmatrix}$$

provide the q-dimensional points that best reconstruct the distances d_{ij} in the original distance matrix.

That is, the Euclidean distance

$$d_E\left(\mathbf{Z}_i, \mathbf{Z}_j\right) \approx d_{ij}$$

Multidimensional Scaling: Classical Solution

If the original distances d_{ij} are the Euclidean distances between the original observations,

$$d_{ij} = d_E(\mathbf{X}_i, \mathbf{X}_j),$$

then the classical MDS solution is the same as the scores on the first q principal components of the original data.

Non-metric Multidimensional Scaling

In contrast to metric MDS, non-metric MDS does not try to match the distances d_{ij} quantitatively, but rather seeks to identify points in low-dimensional space such that the pairwise distances between the low dimensional representation have the same *ordering* as the original pairwise distances/dissimilarities.

That is,

$$d_{ij}^{(q)} > d_{k\ell}^{(q)} \Leftrightarrow d_{ij} > d_{k\ell}$$

Multidimensional Scaling: Non-metric MDS

Non-metric MDS is performed via an iterative algorithm:

1. Obtain the pairwise dissimilarities/distances d_{ij} between the n observations.
2. Start with a trial configuration in q dimensions. That is, choose n q-dimensional points $\mathbf{Z}_1, \ldots, \mathbf{Z}_n$. Determine the distances
$$d_{ij}^{(q)} = d_E(\mathbf{Z}_i, \mathbf{Z}_j)$$
between the items in this q-dimensional representation.

Multidimensional Scaling: Non-metric MDS

3. Find values $\hat{d}_{ij}^{(q)}$ that satisfy the same ordering as the original distances, and minimize the stress:

$$\text{Stress} = \left[\frac{\sum\sum_{i<j}\left(d_{ij}^{(q)2} - \hat{d}_{ij}^{(q)2}\right)^2}{\sum\sum_{i<j} d_{ij}^{(q)4}}\right]$$

   This step is typically done with some monotonic regression method to produce “fitted” distances.

4. Move the q-dimensional representations of the points to achieve distances $d_{ij}^{(q)}$ closer to the $\hat{d}_{ij}^{(q)}$, and then recompute the $\hat{d}_{ij}^{(q)}$

5. Iterate until there is no more improvement in the stress.

Non-metric Multidimensional Scaling

Non-metric MDS is particularly useful when the starting information is not truly a *distance* matrix, but rather a *dissimilarity* matrix, such as might come from user evaluations or qualitative assessments.