MULTIDIMENSIONAL SCALING (MDS)

SETTING: FOR A SET OF N OBJECTS WE HAVE A MEASUREMENT OF DISSIMILARITIES/DISTANCES BETWEEN EACH PAIR OF OBJECTS

δ_{rs} = DISSIMILARITY OF OBJECTS r AND s

e.g. DISTANCES BETWEEN CITIES, DIFFERENCES IN COLORS FROM OPINIONS OF A COLLECTION OF RANDOMLY SELECTED STUDY SUBJECTS, ...

GOAL: FIND A CONFIGURATION OF POINTS IN LOW-DIMENSIONAL SPACE (USUALLY 1,2, OR 3), WITH ONE POINT PER SUBJECT S.T THE DISTANCE/DISSIMILARITY BETWEEN POINTS/SUBJECTS MATCHES ORIGINAL DISSIMILARITIES, i.e.

FIND $\vec{x}_1' \dots \vec{x}_N' \in \mathbb{R}^k$ S.T

$$\|\vec{x}_r' - \vec{x}_s'\| \approx \delta_{rs}$$

* GIVE A LOW-DIMENSIONAL VISUAL REPRESENTATION OF DATA TO HELP VISUALIZE USEFUL INFO. ABOUT RELATIONSHIPS BETWEEN OBJECTS

NOTE: • $\vec{x}_1' \dots \vec{x}_N' \in \mathbb{R}^k$, $\vec{b} \in \mathbb{R}^{k}$, $A \in \mathbb{R}^{k \times k}$ ORTHONORMAL

$$\|(A\vec{x}_r' + \vec{b}) - (A\vec{x}_s + b)\|_2 = \|A(\vec{x}_r - \vec{x}_s)\|_2$$

$$= \sqrt{[A(x_r' - x_s')]^T [A(\vec{x}_r - \vec{x}_s)}$$

$$= \sqrt{(x_r' - x_s')^T \underbrace{A^T A}_{I} (\vec{x}_r' - \vec{x}_s)} = \|\vec{x}_r - \vec{x}_s\|_2$$

* LOW DIMENSIONAL REPRESENTATION IS ONLY UNIQUE UP TO ROTATION/REFLECTION & TRANSLATION USING EUCLIDEAN DISTANCE

CLASSICAL SCALING (YOUNG, HOUSEHOLDER 1930; TORGERSON 1952)

DEF: AN $N \times N$ MATRIX D IS CALLED A DISTANCE MATRIX IF

1) IT IS SYMMETRIC $D_{rs} = D_{sr}$
2) $D_{rr} = 0$
3) $D_{rs} \geq 0 \quad r \neq s$

FURTHERMORE, D IS CALLED EUCLIDEAN IF THERE EXISTS A CONFIGURATION OF POINTS IN SOME EUCLIDEAN SPACE, i.e. $\exists p$ S.T. $\vec{x}_1, \ldots \vec{x}_N \in \mathbb{R}^p$, SUCH THAT

$$
\begin{aligned}
D_{rs}^2 &= (\vec{x}_r - \vec{x}_s)^T (\vec{x}_r - \vec{x}_s) = \|\vec{x}_r - \vec{x}_s\|^2 \\
&= \vec{x}_r^T \vec{x}_r + \vec{x}_s^T \vec{x}_s - 2 \vec{x}_r^T \vec{x}_s \\
&= \|\vec{x}_r\|^2 + \|\vec{x}_s\|^2 - 2 \vec{x}_r^T \vec{x}_s
\end{aligned}
$$

RECOVERING COORDINATES:

GIVEN A EUCLIDEAN DISTANCE MATRIX $D \in \mathbb{R}^{N \times N}$ CAN WE RECOVER THE COORDINATES $\vec{x}_1, \ldots \vec{x}_N$? (ASSUME $\sum_{i=1}^{N} x_{ij} = 0 \quad j = 1, \ldots p$)

YES,

1) USE $D \in \mathbb{R}^{N \times N}$ TO FIND $B \in \mathbb{R}^{N \times N}$, $B_{rs} = \vec{x}_r^T \vec{x}_s$
2) USE $B = XX^T$ TO RECOVER $\vec{x}_1, \ldots \vec{x}_N$ $\qquad X = \begin{bmatrix} \vec{x}_1^T \\ \vdots \\ \vec{x}_N^T \end{bmatrix} \in \mathbb{R}^{N \times p}$

B ← INNER PRODUCT MATRIX

STEP 1: FIND B FROM D

(B ← INNER PRODUCT; D ← DISTANCE)

$$
\begin{aligned}
\text{a)} \quad \frac{1}{N} \sum_{r=1}^{N} D_{rs}^2 &= \frac{1}{N} \sum_{r=1}^{N} \left(\|\vec{x}_r\|^2 + \|\vec{x}_s\|^2 - 2 \vec{x}_r^T \vec{x}_s \right) \\
&= \left(\frac{1}{N} \sum_{r=1}^{N} \|\vec{x}_r\|^2 \right) + \frac{N}{N} \|\vec{x}_s\|^2 - 2 \underbrace{\left(\sum_{r=1}^{N} \vec{x}_r^T \right)}_{=0} \vec{x}_s \\
&= \frac{1}{N} \sum_{r=1}^{N} \|\vec{x}_r\|^2 + \|\vec{x}_s\|^2 = \frac{1}{N} tr(B) + B_{ss}
\end{aligned}
$$

b) $$\frac{1}{N}\sum_{s=1}^{N} d_{rs}^2 = \|\vec{x}_r\|^2 + \frac{1}{N}\sum_{s=1}^{N}\|\vec{x}_s\|^2 = \frac{1}{N}tr(B) + B_{rr}$$

c) $$\frac{1}{N^2}\sum_{r=1}^{N}\sum_{s=1}^{N} D_{rs}^2 = \frac{1}{N^2}\left(\sum_{r=1}^{N}\sum_{s=1}^{N}\|\vec{x}_r\|^2 + \sum_{r=1}^{N}\sum_{s=1}^{N}\|\vec{x}_s\|^2 - 2\left(\sum_{r=1}^{N}\vec{x}_r^T\right)\left(\sum_{s=1}^{N}\vec{x}_s\right)\right)$$

$$= \frac{1}{N^2}\left[2N\sum_{r=1}^{N}\|\vec{x}_r\|^2\right] = \frac{2}{N}\, tr(B)$$

d) $$B_{rs} = \vec{x}_r^T\vec{x}_s = -\frac{1}{2}\left(D_{rs}^2 - \vec{x}_r^T\vec{x}_r - \vec{x}_s^T\vec{x}_s\right)$$

$$= -\frac{1}{2}\left[D_{rs}^2 - \left(\frac{1}{N}\sum_{s=1}^{N}D_{rs}^2 - \frac{1}{N}tr(B)\right) - \left(\frac{1}{N}\sum_{r=1}^{N}D_{rs}^2 - \frac{1}{N}tr(B)\right)\right]$$

$$= -\frac{1}{2}\left[D_{rs}^2 - \frac{1}{N}\sum_{s=1}^{N}D_{rs}^2 - \frac{1}{N}\sum_{r=1}^{N}D_{rs}^2 + \frac{2}{N}tr(B)\right]$$

$$= -\frac{1}{2}\left[\quad '' \qquad '' \qquad '' \qquad \frac{1}{N^2}\sum\sum D_{rs}^2\right]$$

Matrix Notation

Let $a_{rs} = -\frac{1}{2}D_{rs}^2$

$$\left.\begin{aligned} a_{r.} &= \frac{1}{N}\sum_s a_{rs} = -\frac{1}{2N}\sum_s D_{rs}^2 \\ a_{.s} &= \frac{1}{N}\sum_r a_{rs} = -\frac{1}{2N}\sum_s D_{rs}^2 \\ a_{..} &= \frac{1}{N^2}\sum_{r,s} a_{rs} = \frac{1}{N}\sum_{r,s} D_{rs}^2 \end{aligned}\right\} \Rightarrow B_{rs} = a_{rs} - a_{r.} - a_{.s} + a_{..}$$

$$1_N = (1,1,1,\dots,1)^T \in \mathbb{R}^N$$

If $A \in \mathbb{R}^{N\times N}$, $A_{rs} = a_{rs}$ then

$$B = HAH \qquad H = I_N - \frac{1}{N}1_N 1_N^T$$

$$= \left(I - \frac{1}{N}1_N 1_N\right)A\left(I - \frac{1}{N}1_N 1_N\right)$$

$$= A - \underbrace{1_N\left(\frac{1}{N}1_N^T A\right)}_{1_N\begin{pmatrix}a_{1.}\\ \vdots \\ a_{N.}\end{pmatrix}^T} - \underbrace{\left(A\ \frac{1}{N}1_N\right)1_N^T}_{\begin{pmatrix}a_{.1}\\ \vdots \\ a_{.N}\end{pmatrix}1_N^T} + 1_N\underbrace{\left(\frac{1_N^T A 1_N}{N}\right)}_{1_N a_{..} 1_N^T}1_N^T$$

$$\Rightarrow B_{rs} = A_{rs} - a_{r.} - a_{.s} + a_{..}$$

STEP 2: FIND $\vec{x}_1, \ldots \vec{x}_N$ FROM B

RECALL $B_{rs} = \vec{x}_r^T \vec{x}_s \longleftrightarrow B = XX^T$

- WHAT IS RANK (B) ? p

- B IS SYMMETRIC. WHAT IS EIGENDECOMP. OF B ?

$$B = U \Lambda U^T \quad , \quad \Lambda = \begin{bmatrix} \sigma_1 & & & & \\ & \ddots & & & 0 \\ & & \sigma_p & & \\ & & & 0 & \\ & 0 & & & \ddots \\ & & & & & 0 \end{bmatrix} \qquad \tilde{\Lambda} = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_p \end{bmatrix}$$

$$= \tilde{U} \tilde{\Lambda} \tilde{U}^T \qquad \tilde{U} = [\vec{u}_1 | \ldots | \vec{u}_p]$$

- HOW CAN WE FIND X FROM $B = U \Lambda U^T$?

$$B = \underbrace{\tilde{U} \tilde{\Lambda}^{1/2}}_{X} \underbrace{\tilde{\Lambda}^{1/2} \tilde{U}^T}_{X^T} \qquad \Lambda^{1/2} = \begin{bmatrix} \sigma_1^{1/2} & & 0 \\ & \ddots & \\ 0 & & \sigma_p^{1/2} \end{bmatrix}$$

$$\Rightarrow \quad X = [\vec{u}_1 | \ldots | \vec{u}_p] \begin{bmatrix} \sigma_1^{1/2} & & 0 \\ & \ddots & \\ 0 & & \sigma_p^{1/2} \end{bmatrix}$$

NOTE: i) PARITY $\vec{u}_1 \longmapsto -\vec{u}_1$ LEADS TO REFLECTED SOLUTIONS (UNIQUENESS UP TO SIGN)

CLASSICAL SCALING IN PRACTICE

GIVEN A DISTANCE MATRIX D AND WANT TO APPROXIMATELY RETRIEVE IN LOWER DIMENSION.

(NOTE D MAY OR MAY NOT BE EUCLIDEAN)

CLASSICAL SOLUTION TO MDS PROBLEM IN k DIMENSIONS

a) FROM D FORM MATRIX A, $A_{rs} = -\frac{1}{2} D_{rs}^2$

b) OBTAIN $B = HAH$

c) FIND THE k LARGEST EIGENVALUES $\lambda_1 > \ldots \lambda_k$ OF B W/ EIGENVECTORS $\vec{x}_{(1)}, \ldots \vec{x}_{(k)} \in \mathbb{R}^N$

(WE ARE ASSUMING THE FIRST k EIGENVALUES OF B ARE POSITIVE!)

d) MDS COORDINATES OF POINTS ARE ROWS OF $[\vec{x}_{(1)} \ldots \vec{x}_{(k)}] \begin{bmatrix} \lambda_1^{1/2} & & 0 \\ & \ddots & \\ 0 & & \lambda_k^{1/2} \end{bmatrix}$

THM: Let D be a distance matrix and $B = HAH$ w/ $A_{rs} = -\frac{1}{2} D_{rs}^2$

D is Euclidean $\iff$ B is positive semi-definite

Ex: Suppose $D = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$. Is D Euclidean, i.e. does there exist a, $p = 1, 2, \ldots$

s.t. there are points $\vec{x}_1, \ldots \vec{x}_3 \in \mathbb{R}^p$ w/ $\|\vec{x}_i - \vec{x}_j\|_2 = \delta_{ij}$? If

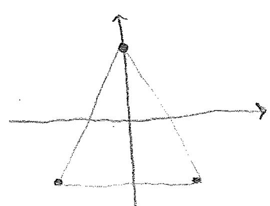
so find the smallest such p and describe the shape formed by

the points $\vec{x}_1, \ldots \vec{x}_3$!

$$A = \begin{bmatrix} 0 & -1/2 & -1/2 \\ -1/2 & 0 & -1/2 \\ -1/2 & -1/2 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 1/3 & -1/6 & -1/6 \\ -1/6 & 1/3 & -1/6 \\ -1/6 & -1/6 & 1/3 \end{bmatrix} = HAH$$

$$= \begin{bmatrix} 0 & \sqrt{2/3} \\ -\sqrt{2}/2 & -\sqrt{1/6} \\ \sqrt{2}/2 & -\sqrt{1/6} \end{bmatrix} \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} 0 & -\sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2/3} & -\sqrt{1/6} & \sqrt{1/6} \end{bmatrix}$$

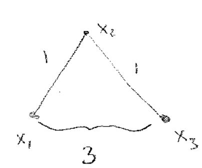
$$X = \begin{bmatrix} 0 & \frac{\sqrt{2}}{\sqrt{3}} \\ -\frac{\sqrt{2}}{2} & -1/\sqrt{6} \\ \frac{\sqrt{2}}{2} & -1/\sqrt{6} \end{bmatrix} \begin{bmatrix} \sqrt{2}/2 & 0 \\ 0 & \sqrt{2}/2 \end{bmatrix} = \begin{bmatrix} 0 & 1/\sqrt{3} \\ -1/2 & -1/2\sqrt{3} \\ +1/2 & -1/2\sqrt{3} \end{bmatrix}$$



Equilateral triangle

* Can rotate about origin and preserve distances *

Ex: $D = \begin{bmatrix} 0 & 1 & 3 \\ 1 & 0 & 1 \\ 3 & 1 & 0 \end{bmatrix}$



$$A = \begin{bmatrix} 0 & -1/2 & -25/2 \\ -1/2 & 0 & -1/2 \\ -25/2 & -1/2 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 17/3 & 7/6 & -41/6 \\ 7/6 & -7/3 & 7/6 \\ -41/6 & 7/6 & 17/3 \end{bmatrix} = \begin{bmatrix} 12.5 & 0 \\ 0 & -3.5 \end{bmatrix}$$

Cannot be configured to match distances

Cannot be negative if Euclidean

HANDLING NEGATIVE EIGENVALUES

OPTION 1: ADD A SUITABLE CONSTANT TO DISSIMILARITIES

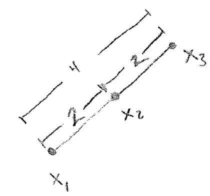
$$D_{rr} = 0$$

$$D_{rs} \to D_{rs} + c \quad \to \quad A \quad \to \quad B \quad \to \quad \lambda_1 \geq \ldots \geq \lambda_k \geq 0$$

IN PREVIOUS EXAMPLE,

$$D' = \begin{bmatrix} 0 & 1 & 3 \\ 1 & 0 & 1 \\ 3 & 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

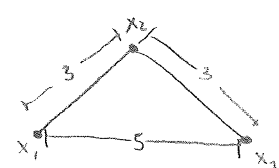
$$= \begin{bmatrix} 0 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 0 \end{bmatrix}$$



* CAN FIND $\overset{\text{EXACT}}{\vee}$ MDS W/ $k=1$! *

$$D'' = \begin{bmatrix} 0 & 1 & 3 \\ 1 & 0 & 1 \\ 3 & 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 3 & 5 \\ 3 & 0 & 3 \\ 5 & 3 & 0 \end{bmatrix}$$



* CAN FIND $\overset{\text{EXACT}}{\vee}$ MDS W/ $k=2$!

OPTION 2. IGNORE <0 EIGENVALUES, IF

$$\lambda_1 > \ldots > \lambda_{k_0} > 0 > \lambda_{k_0+1} > \ldots > \lambda_p$$

THEN TAKE $k \leq k_0$

(SUITABLE WHEN $\lambda_{k_0}, \ldots \lambda_p \approx 0$)

Duality of Classical Scaling and PCA, using Euclidean norm

Suppose we are given $\vec{x}_1, \dots \vec{x}_N \in \mathbb{R}^p$, and asked to find a classical solution to MDS problem

$$\begin{tikzcd}
\vec{x}_1, \dots \vec{x}_N \arrow[r, mapsto] & D \arrow[r, mapsto] & A \arrow[r, mapsto] & B \arrow[dd, leftrightarrow] \\
 & \text{OR} & & \\
X = \begin{bmatrix} \vec{x}_1^T \\ \vdots \\ \vec{x}_N^T \end{bmatrix} \arrow[rrr, mapsto] & & & B = (HX)(HX)^T
\end{tikzcd}$$

$$HX = \begin{bmatrix} \vec{x}_1^T - \bar{x}^T \\ \vdots \\ \vec{x}_N^T - \bar{x}^T \end{bmatrix}$$

Now, $\operatorname{rank}(B) = \operatorname{rank}(HX) = p$

$$\Rightarrow B \text{ has } \underbrace{p \text{ non-negative eigenvalues}}$$

$$\|$$

squares of singular values of HX

$$\|$$

p non-negative eigenvalues of $(X^T H^T H X) = X^T H X = n \hat{\Sigma}_x \qquad (H^2 = H)$

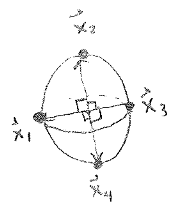
<u>Thm</u>: The MDS coordinates of X in k dimensions are given by the first k principal component scores

* Choosing a suitably small choice of k in classical MDS is equivalent to choosing a suitably small k in PCA

* Classical MDS at dimension k inherits all of the properties of PCA at rank k; e.g. best rank k approx. under Frobenius norm!

<u>Ex</u>: GLOBAL AIRLINE DISTANCES. SEE CANVAS > FILES > NUMERICAL EXAMPLES

<u>Ex</u>: LET $\vec{x}_1, \vec{x}_3$ BE TWO POINTS ON EQUATOR ON OPPOSITE SIDES OF A SPHERE IN $\mathbb{R}^3$ W/ RADIUS $\frac{2}{\pi}$. LET $\vec{x}_2, \vec{x}_4$ BE ON THE NORTH AND SOUTH POLE RESPECTIVELY. LET D_{rs} BE THE LENGTH OF THE SHORTEST PATH FROM $\vec{x}_r$ TO $\vec{x}_s$ ALONG THE SURFACE OF THE SPHERE (GEODESICS)



$$D = \begin{bmatrix} 0 & 1 & 2 & 1 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 1 & 2 & 1 & 0 \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & -1/2 & -2 & -1/2 \\ -1/2 & 0 & -1/2 & -2 \\ -2 & -1/2 & 0 & -1/2 \\ -1/2 & -2 & -1/2 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 3/4 & 1/4 & -5/4 & 1/4 \\ 1/4 & 3/4 & 1/4 & -5/4 \\ -5/4 & 1/4 & 3/4 & 1/4 \\ 1/4 & -5/4 & 1/4 & 3/4 \end{bmatrix}$$

ANGLE BETWEEN $\vec{x}_r, \vec{x}_s$

$*$ IF EUCLIDEAN $\dfrac{\vec{x}_r^T \vec{x}_s}{\|\vec{x}_r\| \cdot \|\vec{x}_s\|} = \cos(\theta_{rs}) = \dfrac{B_{rs}}{\left(\frac{2}{\pi}\right)^2}$

$$\frac{B_{12}}{\left(\frac{2}{\pi}\right)^2} = \frac{\pi^2}{16} \Rightarrow \theta_{rs} = \cos^{-1}\left(\frac{\pi^2}{16}\right) \neq \frac{\pi}{2}$$

Distance Scaling

* Given dissimilarities, δ_{ij} $1 \le i < j \le N$, classical scaling sought a configuration $\vec{y}_1, \ldots \vec{y}_N \in \mathbb{R}^t$ ($t = 1,2,3$ usually) s.t.

$$\|\vec{y}_i - \vec{y}_j\| = d_{ij} \approx \delta_{ij}$$

(viewed dissimilarities as distances)

In distance scaling, we seek a configuration

$$d_{ij} = f(\delta_{ij})$$

monotonic, increasing function
"mapping dissimilarities to distances"

Metric Distance Scaling / Metric MDS	Non-metric Distance Scaling / Nonmetric MDS
• Dissimilarities are quantitative e.g. (Euclidean) distances between points	• Dissimilarities are ordinal e.g. rating scales
• Goal is to preserve "distances" $f(\delta_{ij}) \approx d_{ij}$	• Goal is to preserve rank order $\delta_{ij} < \delta_{k\ell} \Rightarrow f(\delta_{ij}) \le f(\delta_{k\ell})$ $1 \le i,j,k,\ell \le N$

Metric MDS

- δ_{ij} are quantitative measurements using (Euclidean) distance metrics. Alternative metrics include Minkowski $\delta_{ij} = \left(\sum_k (x_{ik} - x_{jk})^p \right)^{1/p}$ $p \ge 1$.

- f is usually taken to be a parametric, monotone function such as

$$f(\delta_{ij}) = \alpha + \beta\,\delta_{ij} \quad , \alpha, \beta \ge 0 \text{ unknown}$$

Absolute MDS $(\alpha = 0, \beta = 1)$ ← Izenman "Not very useful in practice"

Ratio MDS $(\alpha = 0, \beta > 0)$

Interval MDS $(\alpha \ge 0, \beta \ge 0)$

If δ_{ij} are Euclidean this is equivalent to classical scaling

METRIC (LEAST-SQUARES) MDS OR KRUSKAL-SHEPHARD SCALING

- GIVEN A PARAMETRIC FUNCTION FOR f, AND A CONFIGURATION $\vec{y}_1, \ldots \vec{y}_N \in \mathbb{R}^t$ W/ DISTANCE $d_{ij} = \|\vec{y}_i - \vec{y}_j\|$, THEN $\{d_{ij}\}$ CAN BE FIT TO $\{f(\delta_{ij})\}$ THROUGH LEAST SQUARES BY MINIMIZING

METRIC STRESS FUNCTION

$$\text{STRESS}(f, \vec{y}_1, \ldots \vec{y}_N, W) = \sqrt{\sum_{i<j} w_{ij}\left(d_{ij} - f(\delta_{ij})\right)^2}$$

WHERE $W = (w_{ij})^{\geq 0}$ IS A GIVEN MATRIX OF WEIGHTS.

- CHOOSING WEIGHTS

  - Q: SUPPOSE $w_{ij} \geq 0$ ARE FIXED AND $f(\delta_{ij}) = \alpha + \beta\,\delta_{ij}$ W/ $\alpha, \beta \geq 0$ FREE PARAMETERS. WHAT IS MINIMUM OF $\text{STRESS}(\alpha, \beta, \vec{y}_1, \ldots \vec{y}_N, W)$?

    A: SET $\alpha = \beta = 0$, $\vec{y}_1 = \ldots = \vec{y}_N = \vec{0}$ SO $d_{ij} = 0 = f(\delta_{ij})$

    THEN STRESS $= 0$, BUT NOT A MEANINGFUL REPRESENTATION

- SAMMON (NONLINEAR) MAPPING

  - $w_{ij} = \delta_{ij}^{-1} \left\{ \sum_{k<\ell} \delta_{k\ell} \right\}^{-1}$, f = IDENTITY

- KRUSKAL'S (STRESS-1) FORMULA

  - $w_{ij} = \left\{ \sum_{k<\ell} d_{k\ell}^2 \right\}^{-1}$, COMMONLY USED IN NONMETRIC SCALING

- AND MANY OTHERS (SEE IZENMAN TEXT FOR ADDITIONAL DETAILS)

Ex: Sammon v. Classical in Global Cities

Ex. Sammon v Classical in Mixture of Gaussians in $\mathbb{R}^2$

* See Canvas > Files > Numerical Examples > MDS

Observations

- Sammon Mapping

  - $w_{ij} = \underbrace{\delta_{ij}^{-1}}_{\text{Local Normalization}} \underbrace{\left\{ \sum_{k<\ell} \delta_{k\ell} \right\}^{-1}}_{\text{Global Normalization}}$

  - If $\delta_{ij} \ll \delta_{k\ell} \implies w_{ij} \gg w_{k\ell}$

    - Gives greater importance to less dissimilar (more similar) pairs

    - Can be useful for identifying clusters in data

  - f = identity $\Rightarrow$ configuration $\vec{y}_1, \dots, \vec{y}_N$ found by gradient descent

Metric MDS in Practice

1) Fix the number of dimensions t and choose an initial configuration of points (often the classical MDS scaling) and function (parameters) for f

2) Compute d_{ij} and initial stress

3) Change the configuration $\vec{y}_1, \dots, \vec{y}_N$ through gradient descent or pseudo-Newton iterating root finder (Sammon Mapping)

4) Update (parameters) f (optional)

5) Repeat 3 & 4 until convergence

6) Repeat (1) – (5) for various t and compare stress v. t.

NONMETRIC MDS

- SEEKS TO MAINTAIN RANK ORDER OF DISSIMILARITIES
- SEE IZENMAN TEXT FOR DETAILS

Ex: COLOR CIRCLE

* SEE CANVAS > FILES > NUMERICAL EXAMPLES > MDS

PARTING THOUGHTS

- CLASSICAL SCALING APPLIED TO DATA AND EUCLIDEAN DISTANCES "=" PCA
- (NON) METRIC MDS FOCUSES ON MINIMIZING STRESS
  - STRESS = "DEFORMATION" OF DISTANCE IN LOWER DIMENSIONAL REPRESENTATION
  - MANY DIFFERENT FORMS FOR STRESS WHICH PRIORITIZE DIFFERENT RELATIONSHIPS
    - SENSITIVITY TO CHOICE OF STRESS (MEASURE OF DISTANCE)
  - RELIANCE ON NUMERICAL METHODS FOR MINIMIZATION (CAN BE SLOW WHEN N IS LARGE)
  - HOW DOES ONE CHOOSE A MEANINGFUL MEASURE OF DISSIMILARITY/DISTANCE ?