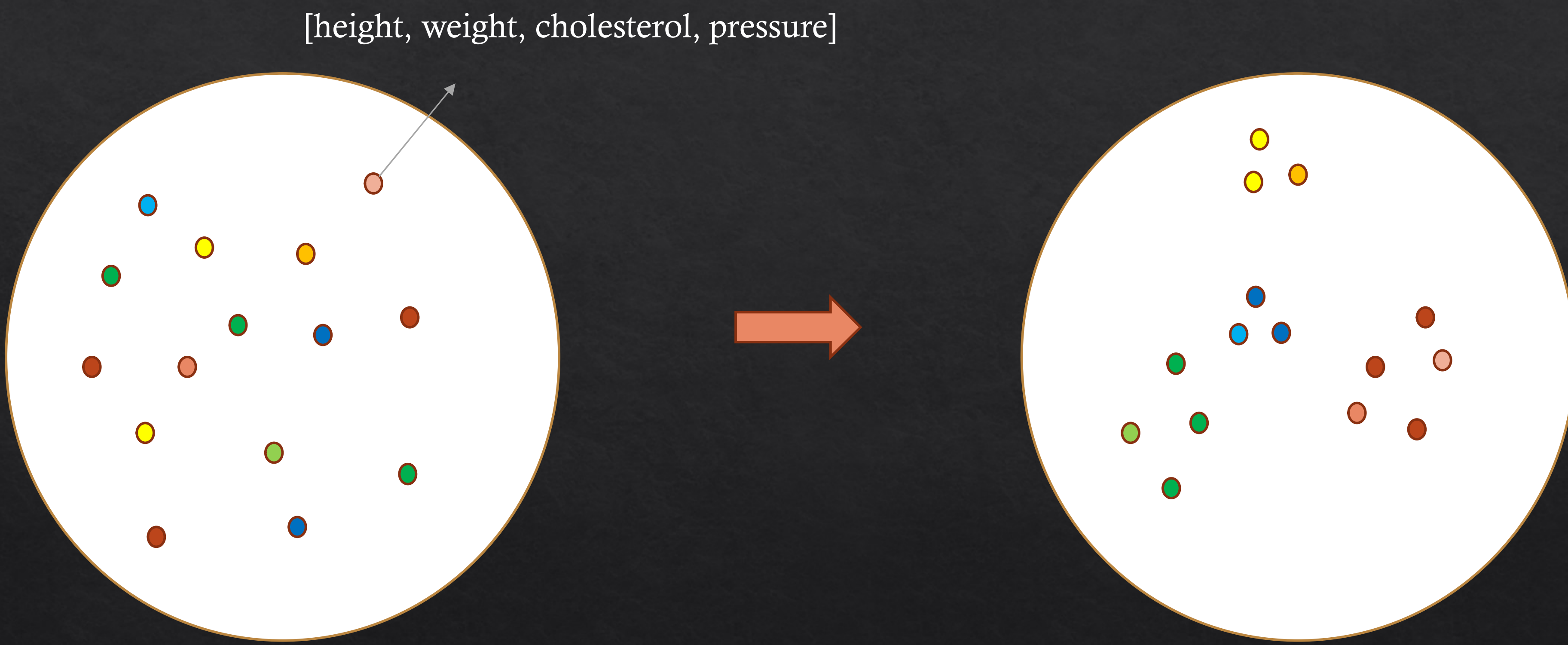
Multi-Dimensional Scaling

Debajyoti Mondal

University of Saskatchewan


[height, weight, cholesterol, pressure]

Given Distances – Find the Locations

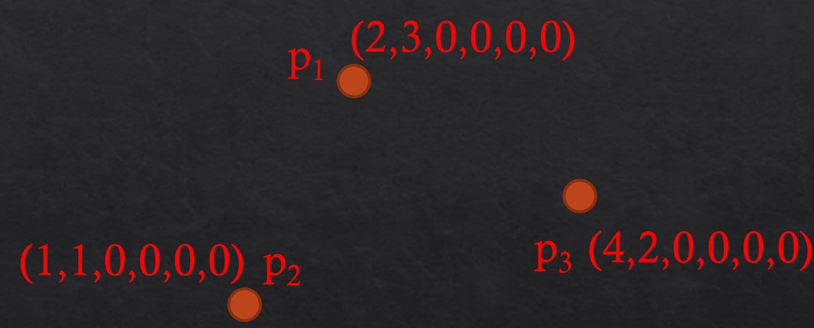
Algorithm

- Set up the matrix of squared proximities $\mathbf{D}^2$
- Apply the double centering: $B = -\frac{1}{2} J \mathbf{D}^2 J$
- Compute m eigenvalues and m eigenvectors
- Multiply eigenvectors and sqrt of eigenvalues

Distance Squared Matrix

$$D^2 = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix}$$

$$D^2 = \begin{bmatrix} 0 & 5 & 5 \\ 5 & 0 & 10 \\ 5 & 10 & 0 \end{bmatrix}$$



$$d_{1,3} = ((2-4)^2 + (3-2)^2) = 5$$

Given Distances – Find the Locations

Algorithm

- ~~Set up the matrix of squared proximities~~ $\cancel{\mathbf{D}^2}$
- Apply the double centering: $B = -\frac{1}{2} J \mathbf{D}^2 J$
- Compute m eigenvalues and m eigenvectors
- Multiply eigenvectors and sqrt of eigenvalues

Double Centering

$$D^2 = \begin{bmatrix} 0 & 5 & 5 \\ 5 & 0 & 10 \\ 5 & 10 & 0 \end{bmatrix}$$

Here 3 comes from the matrix dimension

$$J = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} * -(1/3) \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0.6667 & -0.3333 & -0.3333 \\ -0.3333 & 0.6667 & -0.3333 \\ -0.3333 & -0.3333 & 0.6667 \end{bmatrix}$$

$$B = -\frac{1}{2} J\, D^2\, J = -\frac{1}{2} \begin{bmatrix} 0.6667 & -0.3333 & -0.3333 \\ -0.3333 & 0.6667 & -0.3333 \\ -0.3333 & -0.3333 & 0.6667 \end{bmatrix} * \begin{bmatrix} 0 & 5 & 5 \\ 5 & 0 & 10 \\ 5 & 10 & 0 \end{bmatrix} * \begin{bmatrix} 0.6667 & -0.3333 & -0.3333 \\ -0.3333 & 0.6667 & -0.3333 \\ -0.3333 & -0.3333 & 0.6667 \end{bmatrix}$$

$$B = \begin{bmatrix} 1.1111 & -0.5556 & -0.5556 \\ -0.5556 & 2.7778 & -2.2222 \\ -0.5556 & -2.2222 & 2.7778 \end{bmatrix}$$

Double Centering

$B = -\frac{1}{2} J D^2 J$

Double Centering

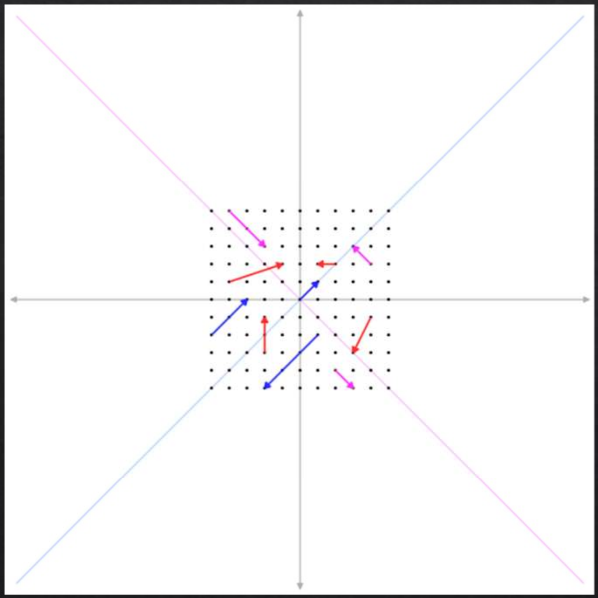
$B = -\frac{1}{2} J D^2 J$

Given Distances – Find the Locations

Algorithm

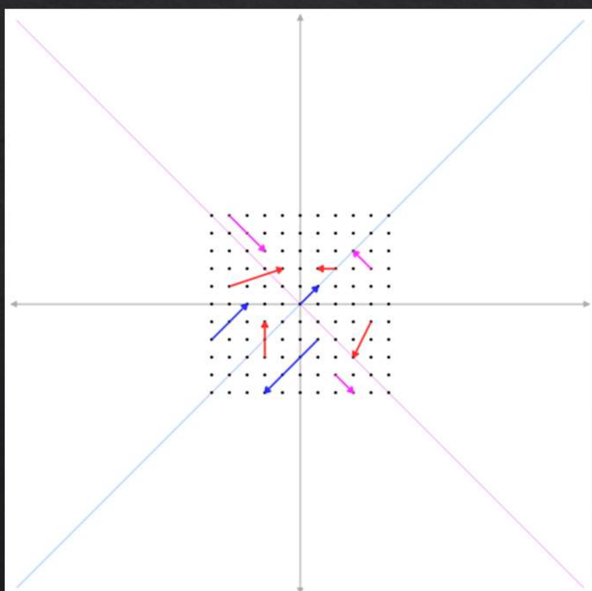
- ~~Set up the matrix of squared proximities~~ $\cancel{\mathbf{D}^2}$
- ~~Apply the double centering~~: $\cancel{B = -\frac{1}{2} J \mathbf{D}^2 J}$
- Compute m eigenvalues and m eigenvectors
- Multiply eigenvectors and sqrt of eigenvalues

Sidetrack: Eigenvector & Eigenvalues



https://en.wikipedia.org/wiki/Eigenvalues_and_eigenvectors

Sidetrack: Eigenvector & Eigenvalues



The transformation matrix $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ preserves the direction of vectors parallel to $v_{\lambda=1} = [1\ {-1}]^T$ (in purple) and $v_{\lambda=3} = [1\ 1]^T$ (in blue). The vectors in red are not parallel to either eigenvector, so, their directions are changed by the transformation. The blue vectors after the transformation are three times the length of the original (their eigenvalue is 3), while the lengths of the purple vectors are unchanged (reflecting an eigenvalue of 1).

https://en.wikipedia.org/wiki/Eigenvalues_and_eigenvectors

Largest 2 Eigenvectors and values

$$B = \begin{bmatrix} 1.1111 & -0.5556 & -0.5556 \\ -0.5556 & 2.7778 & -2.2222 \\ -0.5556 & -2.2222 & 2.7778 \end{bmatrix}$$

$$\text{Eval} = \begin{bmatrix} 0.0000 & 0 & 0 \\ 0 & 1.6667 & 0 \\ 0 & 0 & 5.0000 \end{bmatrix}$$

$$\text{Evec} = \begin{bmatrix} -0.5774 & -0.8165 & -0.0000 \\ -0.5774 & 0.4082 & -0.7071 \\ -0.5774 & 0.4082 & 0.7071 \end{bmatrix}$$

Given Distances – Find the Locations

Algorithm

- ~~Set up the matrix of squared proximities $\mathbf{D}^2$~~
- ~~Apply the double centering: $B = -\frac{1}{2} J \mathbf{D}^2 J$~~
- ~~Compute m eigenvalues and m eigenvectors~~
- Multiply eigenvectors and sqrt of eigenvalues

Largest m Evec and Eval, here m=2

$$B = \begin{bmatrix} 1.1111 & -0.5556 & -0.5556 \\ -0.5556 & 2.7778 & -2.2222 \\ -0.5556 & -2.2222 & 2.7778 \end{bmatrix}$$

$$\text{Eval} = \begin{bmatrix} 0.0000 & 0 & 0 \\ 0 & 1.6667 & 0 \\ 0 & 0 & 5.0000 \end{bmatrix} \qquad \text{Evec} = \begin{bmatrix} -0.5774 & -0.8165 & -0.0000 \\ -0.5774 & 0.4082 & -0.7071 \\ -0.5774 & 0.4082 & 0.7071 \end{bmatrix}$$

$$P = \begin{bmatrix} -0.0000 & -0.8165 \\ -0.7071 & 0.4082 \\ 0.7071 & 0.4082 \end{bmatrix} * \begin{bmatrix} \sqrt{5.0000} & 0 \\ 0 & \sqrt{1.6667} \end{bmatrix}$$

$$= \begin{bmatrix} -0.0000 & -1.0541 \\ -1.5811 & 0.5270 \\ 1.5811 & 0.5270 \end{bmatrix}$$

Final Result!

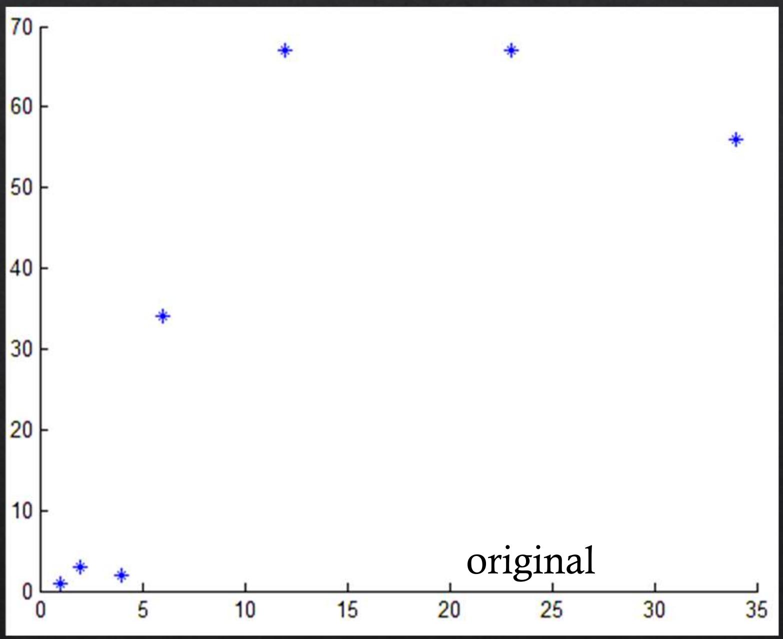
Largest m Evec and Eval, here m=2

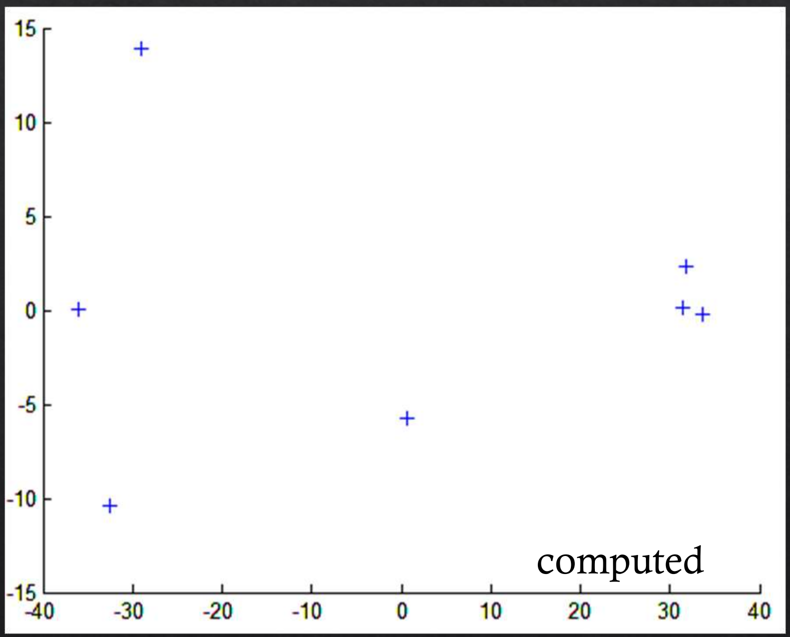
$$B = \begin{bmatrix} 1.1111 & -0.5556 & -0.5556 \\ -0.5556 & 2.7778 & -2.2222 \\ -0.5556 & -2.2222 & 2.7778 \end{bmatrix}$$

Largest m Evec and Eval, here m=2

$$B = \begin{bmatrix} 1.1111 & -0.5556 & -0.5556 \\ -0.5556 & 2.7778 & -2.2222 \\ -0.5556 & -2.2222 & 2.7778 \end{bmatrix}$$

We did not get back the original locations





Given Data → Find Plot (Classic – Multi-Dimensional-Scaling)

Algorithm

- Set up the matrix of squared distances $\mathbf{D}^2$
- Apply the double centering: $B = -\frac{1}{2} J\, \mathbf{D}^2 J$
- Compute m eigenvalues and m eigenvectors
- Multiply eigenvectors and sqrt of eigenvalues

Learning Goal

- Understand dataset types and structural properties of data
- Learn the double centering operation
- Able to relate distance square matrix to scalar product matrix
- Describe the basic idea behind multi-dimensional scaling
- Get familiar with the classic multi-dimensional scaling algorithm