Multidimensional Scaling

Charlie Edelson

10/3/2017

Overview

- Classical Problem
  - Why is it a problem?
  - Principle Coordinate Analysis
  - Relationship to Principle Component Analysis
- General Problem
  - Metric vs Non-metrix

Classical Multidimensional Scaling

A Tale of Three Cities

Imagine we have three cities: A, B, and C

We wish to make a map of our cities, *but* we only know *distances* and not *locations*.

How do we place them on the map?

$\rightarrow$ How do we represent distances?

Distance Matrix

We can introduce the distance matrix D, where, given a measure of distance d,

$$D = \{d_{ij}\} = \{d(\vec{r}_i, \vec{r}_j)\}. \tag{1}$$

For our city example, where $\vec{r}_i = (x_i, y_i)^T$, we'll use

$$d(\vec{r}_B, \vec{r}_A) = \sqrt{(x_b - x_a)^2 + (y_b - y_a)^2}. \tag{2}$$

Example - City Distances

For our cities $d_{ab} = d(\vec{r}_A, \vec{r}_B)$, so

$$D = \begin{bmatrix} d(\vec{r}_A, \vec{r}_A) & d(\vec{r}_A, \vec{r}_B) & d(\vec{r}_A, \vec{r}_C) \\ d(\vec{r}_B, \vec{r}_A) & d(\vec{r}_B, \vec{r}_B) & d(\vec{r}_B, \vec{r}_C) \\ d(\vec{r}_C, \vec{r}_A) & d(\vec{r}_C, \vec{r}_B) & d(\vec{r}_C, \vec{r}_C) \end{bmatrix}. \tag{3}$$

If $\vec{r}_A = (1,1)^T$, $\vec{r}_B = (1,2)^T$, $\vec{r}_C = (2,1)^T$, then

$$D = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & \sqrt{2} \\ 1 & \sqrt{2} & 0 \end{bmatrix}. \tag{4}$$

Rephrase the Problem

Instead of looking for the original vectors $\{\vec{r}_i\}_{i=1}^n$, we instead wish to find the set of vectors $\{\vec{z}_i\}_{i=1}^n$ s.t. they minimize

$$Stress_D = \sum_{i,j} \left(d_{ij} - d(\vec{z}_i, \vec{z}_j) \right). \tag{5}$$

In other words, find a set of vectors which have the same distances as the original vectors.

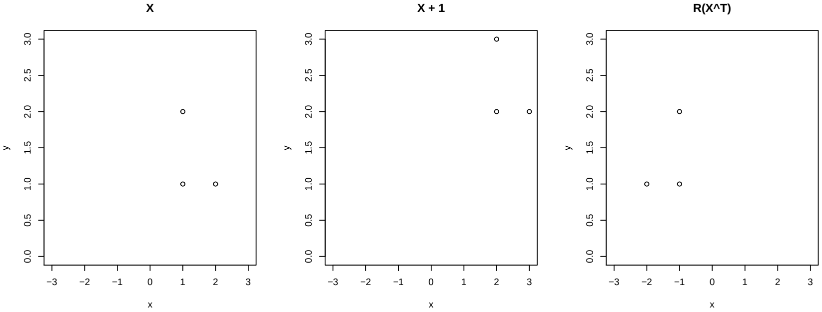
Difficulties Arise

Depending on the definition of $d(\vec{z}_i, \vec{z}_j)$, Eq.8 can difficult to solve analytically.

Additionally...

Unique Selections

...translations and rotations do not affect the distances!



Multiple solutions for a single distance matrix.

Altered Problem

Consider instead the distance matrix B where

$$B = \{b_{ij}\} = \{\langle \vec{x}_i - \bar{x}, \vec{x}_j - \bar{x} \rangle\}. \tag{6}$$

In otherwise B is a matrix of mean centered inner products of $\vec{x}_i$.

Alternatively, for $X \in \mathbb{R}^{n \times p}$ and $X_{i.} = \vec{x}_i - \bar{x}$

$$B = XX^T \tag{7}$$

Altered Problem Cont.

Let review our goal – minimize

$$\sum_{i,j} \left(d_{ij} - d(\vec{z}_i, \vec{z}_j)\right). \tag{8}$$

For B

$$\begin{aligned} Stress &= \sum_{i,j} \left(b_{ij} - d(\vec{z}_i, \vec{z}_j)\right) && (9) \\ &= \sum_{i,j} \left(\langle \vec{x}_i - \bar{x}, \vec{x}_j - \bar{x} \rangle - d(\vec{z}_i, \vec{z}_j)\right) && (10) \end{aligned}$$

Simple Minimum

It becomes evident that given

$$d(\vec{z}_i, \vec{z}_j) = \langle \vec{z}_i, \vec{z}_j \rangle \tag{11}$$

the minimum of

$$Strain_B = \sum_{i,j} \left(\langle \vec{x}_i - \bar{x}, \vec{x}_j - \bar{x} \rangle - d(\vec{z}_i, \vec{z}_j) \right) \tag{12}$$

occures when $\{\vec{z}_i\}_{i=1}^{n} = \{\vec{x}_i - \bar{x}\}$.

Decompose Our Solution

So for $B = XX^T$ we know finding X directly is a solution.

Since B is symetric and semi-definite

$$B = E\Lambda E^{-}1 \tag{13}$$

$$XX^T = E\Lambda^{1/2}\Lambda^{1/2}E^{-1}, \tag{14}$$

E_m is the matrix of eigenvectors of B and Λ is the diagonal matrix of eigenvalues.

Using Eq. 14, we find $X = E\Lambda^{1/2}$.