

# 第11回 多変量解析 多次元尺度法 改訂版

## 多次元尺度法

個体間の類似度が与えられた時に、似ているものは近くに、似ていないものは遠くに、という具合に類似度を距離に変換し、個体間の距離を最も良く再現するような個体の座標を求める方法を多次元尺度法 (Multi-Dimensional Scaling, MDS) と呼ぶ。MDS には、解を解析的に求めようとする古典的多次元尺度法 (Classical MDS) と、データの最適変換を含み、データに距離を直接最小 2 乗法を用いて当てはめる非計量的多次元尺度法 (non-metric MDS) とに大別される。

データとして、個体間の類似度もしくは非類似度が必要。

個体対を提示して評定させる。

個体対を提示して似ているか似ていないかを判断させる。

個体間の移動時間等。

複数の個人からデータを得た場合に、個人差を認めるか否か。

# 距離と座標

$n$  個の個体の  $r$  次元の座標行列  $X$  から計算される個体間の 2 乗距離

$$d_{ij}^2 = \sum_{a=1}^r (x_{ia} - x_{ja})^2, i, j = 1, 2, \dots, n$$

座標  $X$  が与えられた時に距離  $D$  を計算するのは簡単。

MDS はその逆で、距離から座標を出すための方法。

3

## 距離から座標へ： $D$ がユークリッド距離の場合

$$d_{ij}^2 = \sum_{a=1}^r (x_{ia} - x_{ja})^2, i, j = 1, 2, \dots, n$$

$$D^2 = \text{diag}(XX')\mathbf{1}\mathbf{1}' - 2XX' + \mathbf{1}\mathbf{1}'\text{diag}(XX')$$

$\mathbf{1}$  は要素が全て 1 である  $n \times 1$  のベクトル

$$J = I - \frac{1}{n}\mathbf{1}\mathbf{1}'$$

$$J = J' \quad \text{and} \quad J\mathbf{1} = \mathbf{0}, \quad \text{and} \quad \mathbf{1}'J = \mathbf{0}'$$

スカラー積行列

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

$$= -\frac{1}{2}J(\text{diag}(XX')\mathbf{1}\mathbf{1}' - 2XX' + \mathbf{1}\mathbf{1}'\text{diag}(XX'))J$$

$$= JXX'J = YY'$$

中心化された座標行列  $Y$

$$Y = JX$$

$$\text{or} \quad y_{ia} = x_{ia} - \frac{1}{n} \sum_{k=1}^n x_{ka}$$

4

## 距離から座標へ：D がユークリッド距離の場合

$$\begin{aligned} B &= -\frac{1}{2}JD^2J \\ &= -\frac{1}{2}J(\text{diag}(XX')\mathbf{1}\mathbf{1}' - 2XX' + \mathbf{1}\mathbf{1}'\text{diag}(XX'))J \\ &= JXX'J = YY' \end{aligned}$$

$$b_{ij} = -\frac{1}{2} \left( d_{ij}^2 - \frac{1}{n} \sum_{j=1}^n d_{ij}^2 - \frac{1}{n} \sum_{i=1}^n d_{ij}^2 + \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n d_{ij}^2 \right)$$

$$B = P\Lambda P'$$

$$Y = P \left( \Lambda^{\frac{1}{2}} \right)$$

原点  $\mathbf{c}$  と正規直交行列  $\mathbf{T}$  に関して不定性

$$X = YT + \mathbf{1}\mathbf{c}'$$

5

## 古典的 MDS

- データとして  $D$  を得た場合はそれが距離であるとは限らないが  
とりあえずスカラー積  $B$  を計算して最小2乗法で  $Y$  を求める。

$$\text{RSS} = \sum_{i=1}^n \sum_{j=1}^n (b_{ij} - \sum_{a=1}^r y_{ia}y_{ja})^2$$

$$B \approx P_r \Lambda_r P_r'$$

$$Y = P_r \left( \Lambda_r^{\frac{1}{2}} \right)$$

6

# 距離から座標を直接求めることも可能

O をデータとして得られる非類似度行列とする。

$$RSS1 = \sum_{i=2}^n \sum_{j=1}^i (o_{ij} - d_{ij})^2$$

$$d_{ij}^2 = \sum_{a=1}^r (x_{ia} - x_{ja})^2$$

Gauss-Newton 法に必要な Jacobian は

計量的 MDS

$$\frac{\partial d_{kj}}{\partial x_{ka}} = \frac{(x_{ka} - x_{ja})}{d_{kj}}$$

$$\frac{\partial d_{jk}}{\partial x_{ka}} = -\frac{(x_{ka} - x_{ja})}{d_{kj}}$$

Newton-Raphson 法に必要な gradient は

$$\frac{\partial RSS}{\partial x_{ka}} = -2 \sum_{j=1}^n (o_{kj} - d_{kj}) \frac{(x_{ka} - x_{ja})}{d_{kj}} = 0$$

7

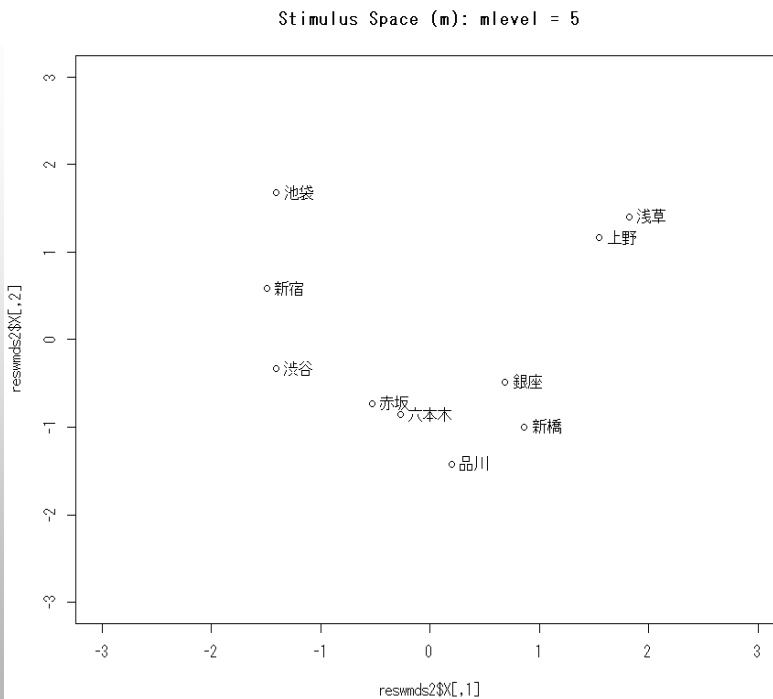
## 数値例

- 21 人による東京の盛り場間の主観的距離（5段階）の平均

	池袋	上野	浅草	銀座	新橋	品川	渋谷	新宿	六本木	赤坂
池袋	0.00	3.63	3.95	3.91	4.47	4.13	2.81	1.61	3.88	3.76
上野	3.63	0.00	1.14	2.82	2.94	3.74	4.31	3.96	3.40	3.70
浅草	3.95	1.14	0.00	3.14	3.45	4.07	4.55	4.43	4.05	3.88
銀座	3.91	2.82	3.14	0.00	1.57	2.05	3.14	3.14	2.00	2.14
新橋	4.47	2.94	3.45	1.57	0.00	1.69	3.12	3.76	2.07	2.29
品川	4.13	3.74	4.07	2.05	1.69	0.00	2.31	3.17	2.12	2.50
渋谷	2.81	4.31	4.55	3.14	3.12	2.31	0.00	1.61	2.29	2.00
新宿	1.61	3.96	4.43	3.14	3.76	3.17	1.61	0.00	2.88	2.64
六本木	3.88	3.40	4.05	2.00	2.07	2.12	2.29	2.88	0.00	1.29
赤坂	3.76	3.70	3.88	2.14	2.29	2.50	2.00	2.64	1.29	0.00

8

# 結果



9

## R code

```
residmds <- function( vecX, 0 ){  
  # rss of metric mds  
  ns=nrow(0)  
  locLH=lower.tri(0, diag=FALSE)  
  X=matrix(vecX, ns)  
  XX=rowSums(X*X)  
  XX11=XX%%matrix(1, 1, ns)  
  D=sqrt( abs( XX11-2*X%%t(X)+t(XX11) ) )  
  # D=distancew(X) [, , 1]  
  resid=c( (0-D)[locLH] )  
  return( resid )  
} # end of residmds
```

$$D^2 = \text{diag}(XX')\mathbf{1}\mathbf{1}' - 2XX' + \mathbf{1}\mathbf{1}'\text{diag}(XX')$$

10

# R code

```
residmds2 <- function( vecX, 0 ){  
  # rss of metric mds  
  ns=nrow(0)  
  locLH=lower.tri(0, diag=FALSE)  
  
  X=matrix(vecX, ns-1)  
  X=rbind(rep(0, ndim), X)  
  
  XX=rowSums(X*X)  
  XX11=XX%%matrix(1, 1, ns)  
  D=sqrt( abs( XX11-2*X%%t(X)+t(XX11) ) )  
  resid=c( (0-D)[locLH] )  
  return( -resid )  
} # end of residmds2
```

原点に関する不定性の除去

パラメタとしては X 行列の  
最初の行を除いたものを用いる。

最初の行は全て 0 とする。

11

## データの最適変換を含む MDS: 非計量的 MDS

### ■ Non-metric MDS

$$RSS1 = \sum_{i=2}^n \sum_{j=1}^i (\tau(o_{ij}) - d_{ij})^2$$

### ■ データの最適変換

- パラメトリックな変換

$$\tau(o) = ao + b, \quad \tau(o) = a \log o + b, \quad \tau(o) = ao^b + c \text{ 等}$$

- 単調増加変換

$$o_{ij} \leq o_{kl} \text{ ならば } \tau(o_{ij}) \leq \tau(o_{kl})$$

### ■ 基準化された RSS

$$RSS1 = \frac{\sum_{i=2}^n \sum_{j=1}^i (\tau(o_{ij}) - d_{ij})^2}{\sum_{i=2}^n \sum_{j=1}^i (\tau(o_{ij}) - \overline{\tau(o_{ij})})^2}$$

12

# 非計量的 MDS の最小2乗解

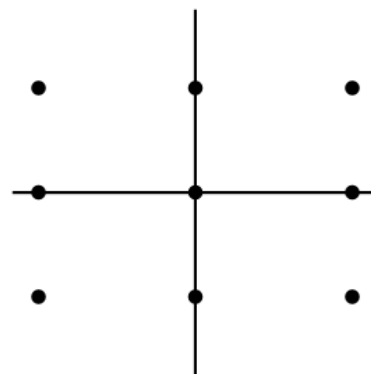
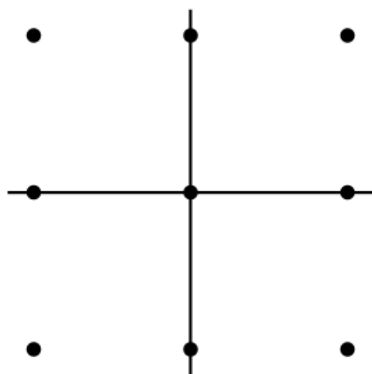
- 母数を、座標行列と最適変換のパラメタに分割して交互最小2乗法を使う。
- 最適変換  $\tau$  が与えられた時の  $X$  は計量的 MDS と同じ。
- 座標  $X$  が与えられた時の最適変換は  
RSS1 を最小にするように求める。
  - ◆ 例えば、線形の場合は  $d$  を  $o$  に回帰すれば良い。
  - ◆  $O$  が順序尺度の場合は最小2乗単調変換を使う。
  - ◆ 基準化された RSS の最小化は Lagrange の未定乗数法を参照。
- 上記のループを実行する自前のプログラムが必要。

13

# 個人差 MDS

- 重み付きユークリッド距離

$$d_{ijk}^2 = \sum_{a=1}^r w_{ka} (x_{ia} - x_{ja})^2$$



$w_{s1} > w_{s2}$  の場合

14

## 最小2乗解

$$\text{RSS1} = \sum_{k=1}^S \sum_{i=2}^n \sum_{j=1}^i (\tau(o_{ijk}) - d_{ijk})^2$$

自由に回転は出来ない。

Jacobian

$$\frac{\partial d_{kjs}}{\partial x_{ka}} = \frac{w_{sa}(x_{ka} - x_{ja})}{d_{kjs}}$$

原点の移動と  
90度ごとの回転と符号の反転のみ。

$$\frac{\partial d_{jks}}{\partial x_{ka}} = -\frac{w_{sa}(x_{ka} - x_{ja})}{d_{kjs}}$$

$$\frac{\partial d_{ijm}}{\partial w_{sa}} = \frac{(x_{ia} - x_{ja})^2}{2d_{ijs}}$$

gradient

$$\frac{\partial \text{RSS}}{\partial x_{ka}} = -2 \sum_{s=1}^S \sum_{j=1}^n (o_{kjs} - d_{kjs}) \frac{w_{sa}(x_{ka} - x_{ja})}{d_{kjs}}$$

$$\frac{\partial \text{RSS}}{\partial w_{sa}} = -\sum_{i>j}^n (o_{ijs} - d_{ijs}) \frac{(x_{ia} - x_{ja})^2}{d_{ijs}}$$

15

## 個人差 MDS の最小2乗解

- 母数を、座標行列、重み行列と最適変換のパラメタに分割して交互最小2乗法を使う。
- 最適変換  $\tau$  と重み  $W$  が与えられた時に  $R_{ss}$  を最小とする  $X$  を求める。
  - ◆ GN を最低1回実行する。(完全な条件付き最小2乗解を求めなくても良い。)
- 最適変換  $\tau$  と座標  $X$  が与えられた時に  $R_{ss}$  を最小とする  $W$  を求める。
  - ◆ GN を最低1回実行する。(完全な条件付き最小2乗解を求めなくても良い。)
- 座標  $X$  と重み  $W$  が与えられた時の最適変換は
  - RSS1 を最小にするように求める。
  - ◆ 例えば、線形の場合は  $d$  を  $o$  に回帰すれば良い。
  - ◆  $O$  が順序尺度の場合は最小2乗単調変換を使う。
  - ◆ 基準化された RSS の最小化は Lagrange の未定乗数法を参照。
- 上記のループを実行する自前のプログラムが必要。

16



# 数値例 21 人分の評定値

0

, , S1

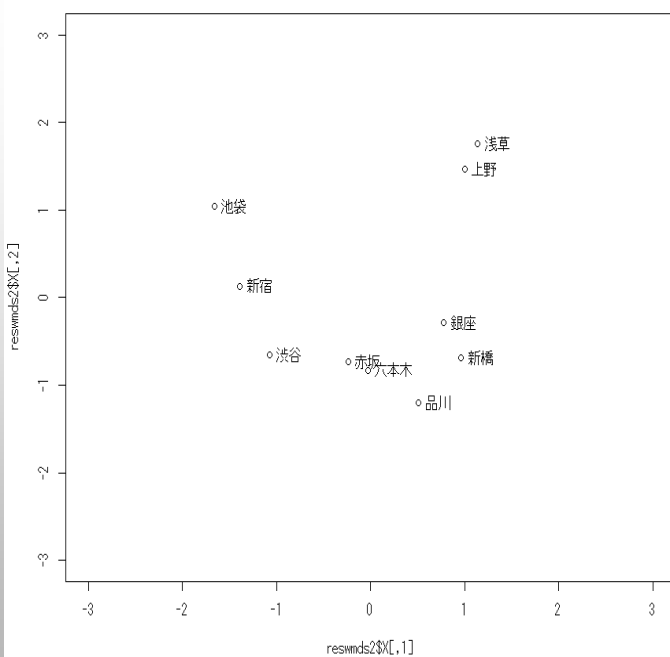
	池袋	上野	浅草	銀座	新橋	品川	渋谷	新宿	六本木	赤坂
池袋	0	5	5	5	5	5	5	2	5	4
上野	5	0	1	2	4	4	5	5	4	4
浅草	5	1	0	4	4	4	5	5	4	4
銀座	5	2	4	0	1	4	4	5	2	4
新橋	5	4	4	1	0	4	5	4	2	2
品川	5	4	4	4	4	0	4	4	4	2
渋谷	5	5	5	4	5	4	0	2	2	2
新宿	2	5	5	5	4	4	2	0	2	2
六本木	5	4	4	2	2	4	2	2	0	1
赤坂	4	4	4	4	2	2	2	2	1	0

, , S2

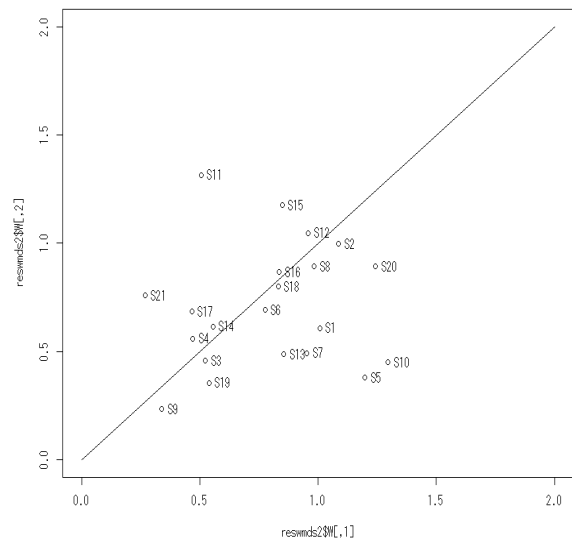
	池袋	上野	浅草	銀座	新橋	品川	渋谷	新宿	六本木	赤坂
池袋	0.0	4	4.0	4.5	5.0	4.5	2.5	2.0	4.5	4
上野	4.0	0	1.0	4.0	3.0	4.0	4.0	4.0	4.0	4
浅草	4.0	1	0.0	3.5	2.0	4.5	4.5	4.0	4.0	4
銀座	4.5	4	3.5	0.0	1.0	2.0	3.5	4.0	1.0	2
新橋	5.0	3	2.0	1.0	0.0	2.0	2.5	4.0	2.0	2
品川	4.5	4	4.5	2.0	2.0	0.0	2.5	3.5	2.0	2
渋谷	2.5	4	4.5	3.5	2.5	2.5	0.0	2.0	2.0	1
新宿	2.0	4	4.0	4.0	4.0	3.5	2.0	0.0	4.0	2
六本木	4.5	4	4.0	1.0	2.0	2.0	2.0	4.0	0.0	1
赤坂	4.0	4	4.0	2.0	2.0	2.0	1.0	2.0	1.0	0

## 線形的最適変換

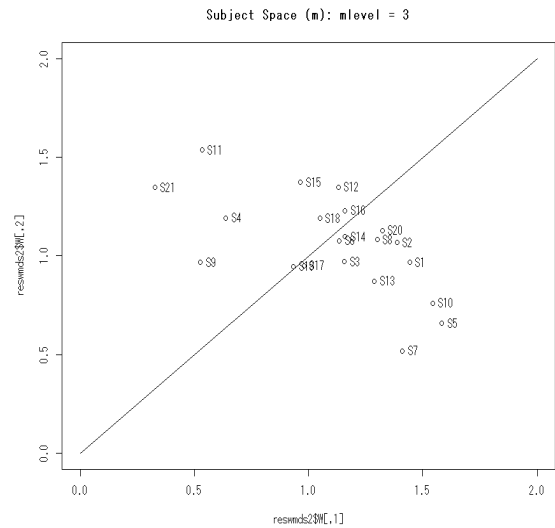
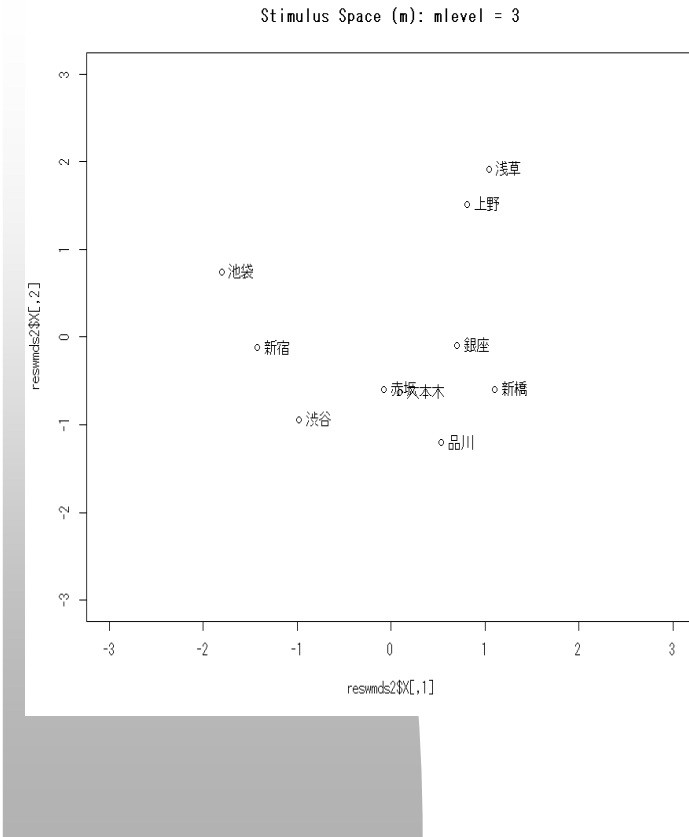
Stimulus Space (m): mlevel = 5



Subject Space (m): mlevel = 5



# 最小2乗単調変換



## 計数データへの拡張

$o_{ij}$  を個体  $i$  と  $j$  が似ていないと判断された場合に 1 となるデータとする。

$D_{ij}$  を個体間の心理的距離を表す確率変数として以下の分布を考える。

$$D_{ij} \sim N(d_{ij}, \sigma^2) \quad (149)$$

$D_{ij}$  がある閾値を超える場合に似ていないという反応が起きる。

$$\Pr(O_{ij} = 1 | d_{ij}, \sigma) = \Pr(D_{ij} > \gamma) \quad (150)$$

$$= \Phi\left(\frac{d_{ij} - \gamma}{\sigma}\right) \approx \Psi\left(\frac{d_{ij} - \gamma}{\sigma}\right) \quad (151)$$

$f_{ij}$  を  $n_{ij}$  回類似度を聞いたときに似ていないと答えた回数とすると尤度関数は

$$L = \prod_{i>j} \Psi\left(\frac{d_{ij} - \gamma}{\sigma}\right)^{f_{ij}} \left(1 - \Psi\left(\frac{d_{ij} - \gamma}{\sigma}\right)\right)^{n_{ij} - f_{ij}} \quad (152)$$