

다차원 척도법(Multi-dimensional Scaling)

여러 대상의 특징 사이 관계에 대한 수치적 자료를 이용하여 유사성에 대한 측정치를 상대적 거리로 구조화하는 방법

2차원 또는 3차원에서의 특정 위치에 관측치를 배치해서 보기 쉽게 척도화

즉, 항목 사이 거리를 기준으로 하는 자료를 이용하여 항목들의 상대적인 위치를 찾고 거리가 가까운 개체들끼리 Group 화 하여 분류할 수 있다.

다차원 척도법 적용 절차

- 1) 자료 수집: 특성을 측정
- 2) 유사성, 비유사성 측정: 개체 사이의 거리 측정
- 3) 공간에서 개체 사이 거리 표현
- 4) 개체의 상호 위치에 따른 관계가 개체들 사이 비유사성에 적합여부 결정

다차원 척도법의 종류

- 1) 계량적(전통적) 다차원 척도법(Classical MDS)

숫자 데이터로만 구성.

stats패키지의 cmdscale() 함수

<https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/cmdscale>

실습

eurodist 데이터를 이용하여 유럽 주요 도시 사이의 거리 대상으로 다차원 척도법을 적용하여 그래프로 표현

=====

```
install.packages("MASS")
```

```
library(MASS)
```

```
data("eurodist")
```

```
eurodist
```

```
# 다차원 척도법 적용
```

```
MDSeurodist <- cmdscale(eurodist)
```

MDSeurodist

시각화

plot(MDSeurodist)

text(MDSeurodist, rownames(MDSeurodist), cex=0.8, col="blue")

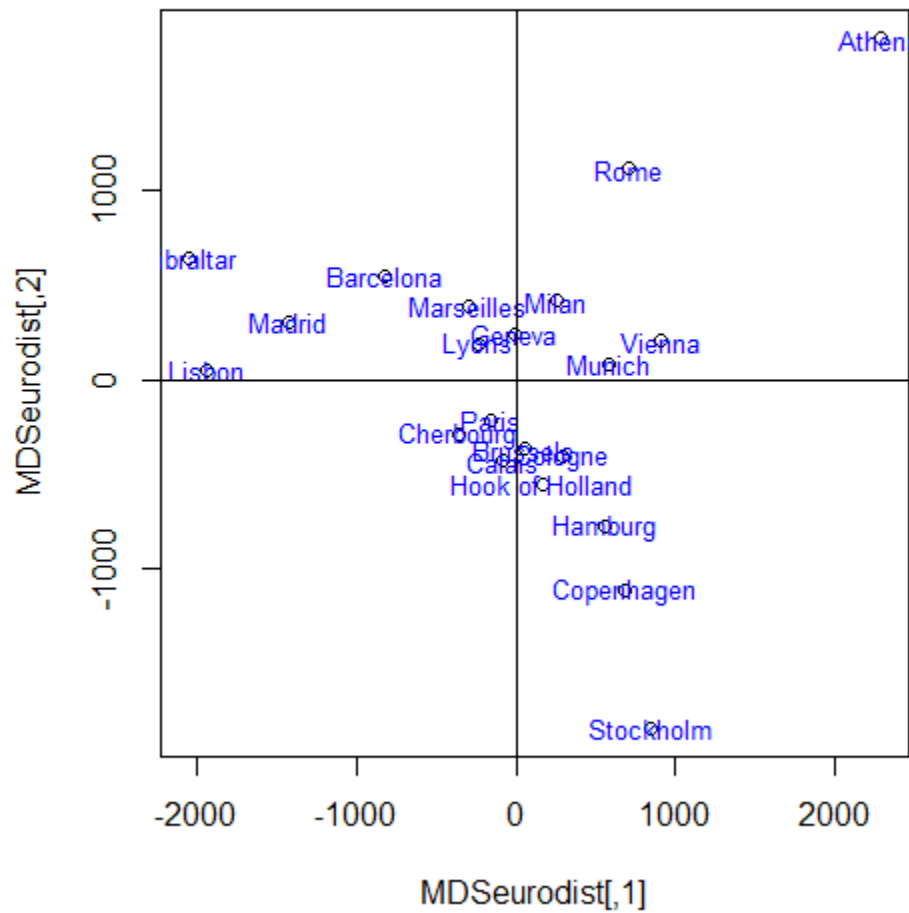
abline(v=0, h=0, lty=1, lwd=0.5)

=====

```
> MDSeurodist <- cmdscale(eurodist)
> MDSeurodist
```

	[,1]	[,2]
Athens	2290.274680	1798.80293
Barcelona	-825.382790	546.81148
Brussels	59.183341	-367.08135
Calais	-82.845973	-429.91466
Cherbourg	-352.499435	-290.90843
Cologne	293.689633	-405.31194
Copenhagen	681.931545	-1108.64478
Geneva	-9.423364	240.40600
Gibraltar	-2048.449113	642.45854
Hamburg	561.108970	-773.36929
Hook of Holland	164.921799	-549.36704
Lisbon	-1935.040811	49.12514
Lyons	-226.423236	187.08779
Madrid	-1423.353697	305.87513
Marseilles	-299.498710	388.80726
Milan	260.878046	416.67381
Munich	587.675679	81.18224
Paris	-156.836257	-211.13911
Rome	709.413282	1109.36665
Stockholm	839.445911	-1836.79055
Vienna	911.230500	205.93020

```
> plot(MDSeurodist)
> text(MDSeurodist, rownames(MDSeurodist), cex=0.8, col="blue")
> abline(v=0, h=0, lty=1, lwd=0.5)
> |
```



그래프를 통해 Paris, Munich 등이 중심에 있고 Lisbon, Stockholm, Athen 등은 중심에서 거리가 좀 있음을 알 수 있다.

이 그래프를 통해서 거리가 가까운 도시끼리 Group화가 가능하다.

2) 비계량적 다차원 척도법 (nonmetric MDS):

숫자가 아닌 데이터 포함.

MASS패키지의 isoMDS() 함수 이용

<https://www.rdocumentation.org/packages/MASS/versions/7.3-53.1/topics/isoMDS>

실습.

HSAUR 패키지 내 voting 데이터를 이용

15명의 의원이 19개의 법안에 투표한 결과 데이터

```
=====
```

```
install.packages("HSAUR")
```

```
library(HSAUR)
```

```
library(MASS)
```

```
data("voting", package="HSAUR")
```

```
voting
```

```
MDS2voting <- isoMDS(voting)
```

```
MDS2voting
```

```
x <- MDS2voting$point[,1]
```

```
y <- MDS2voting$point[,2]
```

```
plot(x,y)
```

```
text(x, y, labels= colnames(voting))
```

```
=====
```

```

> data("voting", package="HSAUR")
> voting

```

	Hunt(R)	Sandman(R)	Howard(D)	Thompson(D)	Freylinghuysen(R)
Hunt(R)	0	8	15	15	10
Sandman(R)	8	0	17	12	13
Howard(D)	15	17	0	9	16
Thompson(D)	15	12	9	0	14
Freylinghuysen(R)	10	13	16	14	0
Forsythe(R)	9	13	12	12	8
widnall(R)	7	12	15	13	9
Roe(D)	15	16	5	10	13
Heltoski(D)	16	17	5	8	14
Rodino(D)	14	15	6	8	12
Minish(D)	15	16	5	8	12
Rinaldo(R)	16	17	4	6	12
Maraziti(R)	7	13	11	15	10
Daniels(D)	11	12	10	10	11
Patten(D)	13	16	7	7	11

	Forsythe(R)	widnall(R)	Roe(D)	Heltoski(D)	Rodino(D)
Hunt(R)	9	7	15	16	14
Sandman(R)	13	12	16	17	15
Howard(D)	12	15	5	5	6
Thompson(D)	12	13	10	8	8
Freylinghuysen(R)	8	9	13	14	12
Forsythe(R)	0	7	12	11	10
widnall(R)	7	0	17	16	15
Roe(D)	12	17	0	4	5
Heltoski(D)	11	16	4	0	3
Rodino(D)	10	15	5	3	0
Minish(D)	9	14	5	2	1
Rinaldo(R)	10	15	3	1	2
Maraziti(R)	6	10	12	13	11
Daniels(D)	6	11	7	7	4
Patten(D)	10	13	6	5	6

	Minish(D)	Rinaldo(R)	Maraziti(R)	Daniels(D)	Patten(D)
Hunt(R)	15	16	7	11	13
Sandman(R)	16	17	13	12	16
Howard(D)	5	4	11	10	7
Thompson(D)	8	6	15	10	7
Freylinghuysen(R)	12	12	10	11	11
Forsythe(R)	9	10	6	6	10
widnall(R)	14	15	10	11	13
Roe(D)	5	3	12	7	6
Heltoski(D)	2	1	13	7	5
Rodino(D)	1	2	11	4	6
Minish(D)	0	1	12	5	5
Rinaldo(R)	1	0	12	6	4
Maraziti(R)	12	12	0	9	13
Daniels(D)	5	6	9	0	9
Patten(D)	5	4	13	9	0

```

> |

```

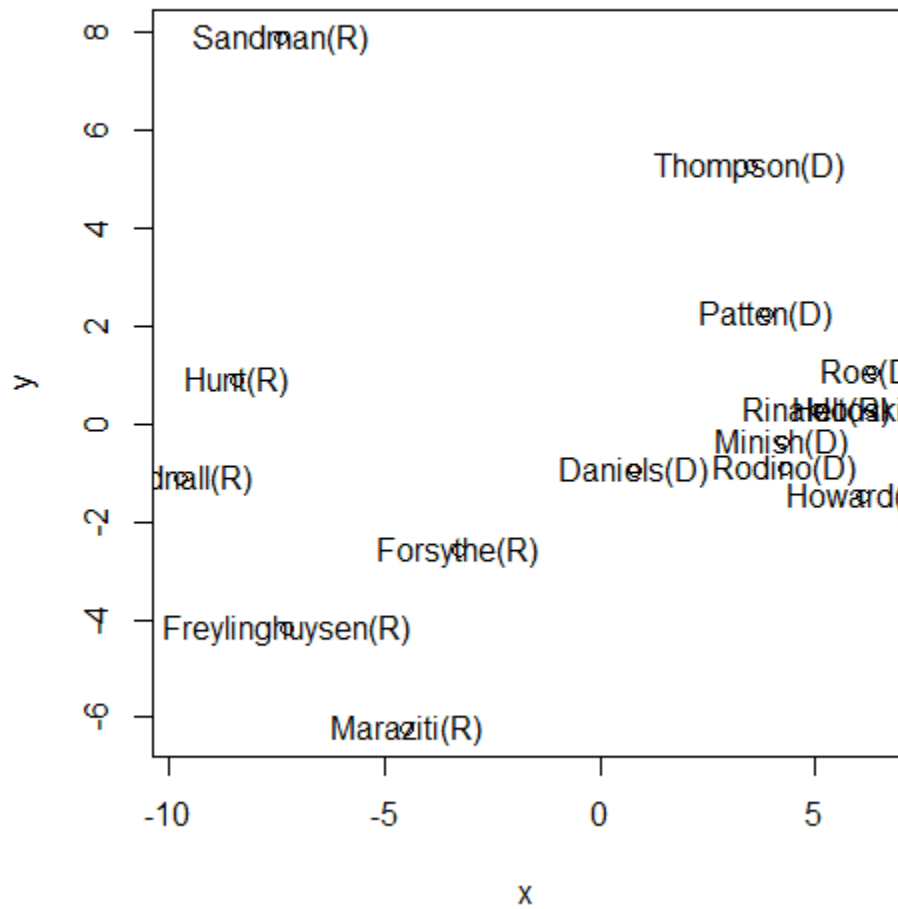
```

> MDS2voting <- isoMDS(voting)
initial value 15.268246
iter 5 value 10.264075
final value 9.879047
converged
> MDS2voting
$points
      [,1]      [,2]
Hunt(R) -8.4354008  0.9063380
Sandman(R) -7.4050250  7.8770232
Howard(D)  6.0930164 -1.4971986
Thompson(D)  3.5187022  5.2486888
Freylinghuysen(R) -7.2457425 -4.1821704
Forsythe(R) -3.2787096 -2.5689673
Widhall(R) -9.7110008 -1.1187710
Roe(D)  6.3429759  1.0388694
Heltoski(D)  6.2983842  0.2706499
Rodino(D)  4.2829160 -0.9151604
Minish(D)  4.2642545 -0.3919690
Rinaldo(R)  5.0285425  0.2665701
Maraziti(R) -4.4577693 -6.2177727
Daniels(D)  0.8129854 -0.9417672
Patten(D)  3.8918709  2.2256372

$stress
[1] 9.879047

> |

```



그래프를 통해 유사한 성향의 의원을 파악하기 쉽다.

다차원 척도법은 주어진 데이터를 기반으로 수행하는 일종의 군집 분석이라고 할 수 있다.