

DƯƠNG MINH LƯƠNG 18521071

EUCLID

Index	0	1	2	3	4	5	6	7	8	9
0	0	60.688	42.733	44.151	53.286	15.27	22.861	29.283	66.167	53.223
1	60.688	0	80.745	64.647	39.564	72.344	45.82	70.144	119.835	99.587
2	42.733	80.745	0	20.306	49.765	52.716	37.285	13.815	43.582	19.98
3	44.151	64.647	20.306	0	30.031	57.925	28.134	20.176	63.888	40.093
4	53.286	39.564	49.765	30.031	0	68.515	30.428	44.265	92.991	69.744
5	15.27	72.344	52.716	57.925	68.515	0	38.107	40.598	66.078	58.561
6	22.861	45.82	37.285	28.134	30.428	38.107	0	24.94	74.029	54.6
7	29.283	70.144	13.815	20.176	44.265	40.598	24.94	0	50.209	29.667
8	66.167	119.835	43.582	63.888	92.991	66.078	74.029	50.209	0	24.24
9	53.223	99.587	19.98	40.093	69.744	58.561	54.6	29.667	24.24	0

Nhận xét : + Ma trận có đường chéo chính là 0

+ Index(I,j) đối với Index(j,i) qua đường chéo chính bằng nhau

+ Khoảng cách giữa các điểm trong không gian N chiều là 1 ma trận , khoảng cách giữa 2 điểm trong mặt phẳng OXY

+ Khoảng cách giữa các điểm tương đối lớn

+ Ma trận vuông với 10x10 với 10 điểm dữ liệu

+ Khoảng vật lý

```
import pandas as pd
import numpy as np
import math
from numpy.linalg import norm
dataset =pd.read_csv("Point_Data.csv")
p=dataset.iloc[0,[0,1]]
q=dataset.iloc[1,[0,1]]
Eculid=norm(p-q)
print("Khoảng cách Eculid = ",Eculid)
K=[]
for i in dataset.itertuples():
    K.append((round(i.x,3),round(i.y,3)))

re=[]
for i in range(10):
```

```

x1,y1=K[i]
row=[]
p=np.array([x1,y1])
for j in range(10):
    x2,y2=K[j]
    q=np.array([x2,y2])
    dist = norm(p-q)
    row.append(round(dist,3))
re.append(row)

dfs = pd.DataFrame(re)
dfs.to_csv("resultEculid10.csv",index=False)

```

Chessboard

Index	0	1	2	3	4	5	6	7	8	9
0	0	55.556	41.721	34.085	51.162	14.259	21.875	27.912	58.46	52.699
1	55.556	0	66.146	58.51	39.319	69.815	33.681	52.337	86.548	77.124
2	41.721	66.146	0	18.816	41.915	47.185	35.081	13.809	40.239	16.694
3	34.085	58.51	18.816	0	23.099	42.322	27.445	19.208	59.055	35.51
4	51.162	39.319	41.915	23.099	0	65.421	29.287	42.307	82.154	58.609
5	14.259	69.815	47.185	42.322	65.421	0	36.134	33.376	63.924	58.163
6	21.875	33.681	35.081	27.445	29.287	36.134	0	21.272	52.867	46.059
7	27.912	52.337	13.809	19.208	42.307	33.376	21.272	0	39.847	24.787
8	58.46	86.548	40.239	59.055	82.154	63.924	52.867	39.847	0	23.545
9	52.699	77.124	16.694	35.51	58.609	58.163	46.059	24.787	23.545	0

Nhận xét : + Ma trận có đường chéo chính là 0

- + Index(I,j) đối với Index(j,i) qua đường chéo chính bằng nhau
- + Các phần tử là kết quả Max giữa xi-yi điểm khi trừ nhau
- + Khoảng cách giữa các điểm tương đối lớn
- + Ma trận vuông với 10x10 với 10 điểm dữ liệu

```

import pandas as pd
import numpy as np
from scipy.spatial import distance
dataset =pd.read_csv("Point_Data.csv")

```

```

K=[]
for i in dataset.itertuples():
    K.append((round(i.x,3),round(i.y,3)))
re=[]
for i in range(10):
    x1,y1=K[i]
    row=[]
    p=np.array([x1,y1])
    for j in range(10):
        x2,y2=K[j]
        q=np.array([x2,y2])
        dist = distance.chebyshev(p,q)
        row.append(round(dist,3))
    re.append(row)

dfs = pd.DataFrame(re)
dfs.to_csv("resultChessboard10.csv",index=False)

```

Hamming

Index	0	1	2	3	4	5	6	7	8	9
0	0	1	1	1	1	1	1	1	1	1
1	1	0	1	1	1	1	1	1	1	1
2	1	1	0	1	1	1	1	1	1	1
3	1	1	1	0	1	1	1	1	1	1
4	1	1	1	1	0	1	1	1	1	1
5	1	1	1	1	1	0	1	1	1	1
6	1	1	1	1	1	1	0	1	1	1
7	1	1	1	1	1	1	1	0	1	1
8	1	1	1	1	1	1	1	1	0	1
9	1	1	1	1	1	1	1	1	1	0

Nhận xét : + Ma trận có đường chéo chính là 0

+ Index(I,j) đối với Index(j,i) qua đường chéo chính bằng nhau

+Index(I,j) bằng 1 không thuộc đường chéo chính vì các số đa số khác nhau

+ Ma trận vuông với 10x10 với 10 điểm dữ liệu

+ Khoảng cách về sự khác nhau giữa 2 điểm

```

import pandas as pd
import numpy as np
import math
from numpy.linalg import norm
from scipy.spatial import distance
from scipy.spatial.distance import pdist

dataset =pd.read_csv("Point_Data.csv")
K=[]
for i in dataset.itertuples():
    K.append((round(i.x,3),round(i.y,3)))
re=[]
for i in range(10):
    x1,y1=K[i]
    row=[]
    p=np.array([x1,y1])
    for j in range(10):
        x2,y2=K[j]
        q=np.array([x2,y2])
        dist = distance.hamming(p,q)
        row.append(round(dist,3))
    re.append(row)

dfs = pd.DataFrame(re)
dfs.to_csv("result Hamming10.csv",index=False)

```

Mahattan

Index	0	1	2	3	4	5	6	7	8	9
0	0	79.981	50.968	62.148	66.056	19.723	28.515	36.767	89.452	60.146
1	79.981	0	112.455	86.003	43.713	88.776	64.746	99.038	169.433	140.127
2	50.968	112.455	0	26.452	68.742	70.691	47.709	14.201	56.978	27.672
3	62.148	86.003	26.452	0	42.29	81.871	33.633	25.381	83.43	54.124
4	66.056	43.713	68.742	42.29	0	85.779	37.541	55.325	125.72	96.414
5	19.723	88.776	70.691	81.871	85.779	0	48.238	56.49	80.657	64.975
6	28.515	64.746	47.709	33.633	37.541	48.238	0	34.292	104.687	75.381
7	36.767	99.038	14.201	25.381	55.325	56.49	34.292	0	70.395	41.089
8	89.452	169.433	56.978	83.43	125.72	80.657	104.687	70.395	0	29.306
9	60.146	140.127	27.672	54.124	96.414	64.975	75.381	41.089	29.306	0

Nhận xét : + Ma trận có đường chéo chính là 0

- + Index(I,j) đối với Index(j,i) qua đường chéo chính bằng nhau**
- + Ma trận vuông với 10x10 với 10 điểm dữ liệu**
- + Khoảng cách giữa các điểm tương đối lớn có thể là lớn nhất**

```
import pandas as pd
import numpy as np
from scipy.spatial import distance

dataset =pd.read_csv("Point_Data.csv")
K=[]
for i in dataset.itertuples():
    K.append((round(i.x,3),round(i.y,3)))
re=[]
for i in range(10):
    x1,y1=K[i]
    row=[]
    p=np.array([x1,y1])
    for j in range(10):
        x2,y2=K[j]
        q=np.array([x2,y2])
        dist = distance.cityblock(p,q)
        row.append(round(dist,3))
    re.append(row)

dfs = pd.DataFrame(re)
dfs.to_csv("resultMahattan10.csv",index=False)
```

Minkowski

Index	0	1	2	3	4	5	6	7	8	9
0	0	57.087	41.872	39.515	51.579	14.522	22.077	28.206	61.23	52.749
1	57.087	0	72.981	60.467	39.337	70.278	40.854	62.591	106.786	89.162
2	41.872	72.981	0	19.226	45.298	49.054	35.618	13.809	41.182	18.146
3	39.515	60.467	19.226	0	26.867	51.635	27.549	19.418	60.408	37.139
4	51.579	39.337	45.298	26.867	0	66.072	29.504	42.714	86.05	63.443
5	14.522	70.278	49.054	51.635	66.072	0	36.581	36.724	64.304	58.194
6	22.077	40.854	35.618	27.549	29.504	36.581	0	22.787	65.955	49.721
7	28.206	62.591	13.809	19.418	42.714	36.724	22.787	0	45.107	26.944
8	61.23	106.786	41.182	60.408	86.05	64.304	65.955	45.107	0	23.659
9	52.749	89.162	18.146	37.139	63.443	58.194	49.721	26.944	23.659	0

Nhận xét : + Ma trận có đường chéo chính là 0

+ Ma trận vuông với 10x10 với 10 điểm dữ liệu

+ Index(i,j) đối với Index(j,i) qua đường chéo chính bằng nhau

+ Nếu r =1 thì là Mahattan

+ Nếu r = 2 thì là Euclid

Ở đây thì dùng r = 3

```
import pandas as pd
import numpy as np
import math
from numpy.linalg import norm
from scipy.spatial import distance
from scipy.spatial.distance import pdist

dataset =pd.read_csv("Point_Data.csv")
K=[]
for i in dataset.itertuples():
    K.append((round(i.x,3),round(i.y,3)))
re=[]
r=int(input("Nhập r:"))
for i in range(10):
    x1,y1=K[i]
    row=[]
    p=np.array([x1,y1])
    for j in range(10):
```

```

x2,y2=K[j]
q=np.array([x2,y2])
dist = distance.minkowski(p,q,r)
row.append(round(dist,3))
re.append(row)

dfs = pd.DataFrame(re)
dfs.to_csv("resultminkowski10.csv",index=False)

```

Canberra

Index	0	1	2	3	4	5	6	7	8	9
0	0	0.581	0.604	0.639	0.533	0.268	0.277	0.388	1.54	0.828
1	0.581	0	0.936	0.675	0.315	0.711	0.434	0.758	1.755	1.294
2	0.604	0.936	0	0.317	0.693	0.853	0.57	0.25	1.381	0.562
3	0.639	0.675	0.317	0	0.398	0.873	0.372	0.265	1.53	0.841
4	0.533	0.315	0.693	0.398	0	0.749	0.274	0.465	1.676	1.135
5	0.268	0.711	0.853	0.873	0.749	0	0.523	0.645	1.403	0.864
6	0.277	0.434	0.57	0.372	0.274	0.523	0	0.354	1.632	1.014
7	0.388	0.758	0.25	0.265	0.465	0.645	0.354	0	1.5	0.757
8	1.54	1.755	1.381	1.53	1.676	1.403	1.632	1.5	0	1.004
9	0.828	1.294	0.562	0.841	1.135	0.864	1.014	0.757	1.004	0

Nhận xét : + Ma trận có đường chéo chính là 0

+ Index(i,j) đối với Index(j,i) qua đường chéo chính bằng nhau

+ Ma trận vuông với 10x10 với 10 điểm dữ liệu

```
import pandas as pd
import numpy as np
import random
from scipy.spatial import distance

dataset =pd.read_csv("Point_Data.csv")
K=[]
for i in dataset.itertuples():
    K.append((round(i.x,3),round(i.y,3)))
re=[]
for i in range(10):
    x1,y1=K[i]
    row=[]
    p=np.array([x1,y1])
    for j in range(10):
        x2,y2=K[j]
        q=np.array([x2,y2])
        dist = distance.canberra(p,q)
        row.append(round(dist,3))
    re.append(row)

dfs = pd.DataFrame(re)
dfs.to_csv("resultCanberra10.csv",index=False)
```

Stdeuclid

[illegible]

Nhận xét : + Ma trận có đường chéo chính là 0

+ Index(i,j) đối với Index(j,i) qua đường chéo chính bằng nhau

+ Không xác định được kết quả

+ Ma trận vuông với 10x10 với 10 điểm dữ liệu

```
import pandas as pd
import numpy as np
from scipy.spatial.distance import pdist
import random
dataset = pd.read_csv("Point_Data.csv")
K=[]
for i in dataset.itertuples():
    K.append((round(i.x,3),round(i.y,3)))
re=[]
for i in range(10):
    x1,y1=K[i]
    row=[]
    p=np.array([x1,y1])
    for j in range(10):
        x2,y2=K[j]
        q=np.array([x2,y2])
        dist = pdist([p,q], 'seuclidean', V=None)
        row.append(dist)
    re.append(row)

dfs = pd.DataFrame(re)
dfs.to_csv("resultstdeuclid10.csv",index=False)
```