9/25

ISYE 6740 Homework 2

Q1

1. Show that log-sum-exp $f(x) = \log(e^{x_1} + \dots + e^{x_n})$ is convex

Proof: Let $u_i = e^{x_i}$ $v_i = e^{y_i}$

$$\Rightarrow \underbrace{f(\theta x + (1-\theta)y)}_{\mathrm{I}} = \log\left(\sum_{i=1}^{n} e^{\theta x_i + (1-\theta)y_i}\right) = \log\left(\sum_{i=1}^{n} u_i^{\theta} v_i^{(1-\theta)}\right)$$

Applying Hölders inequality with:

$$\sum_{i=1}^{n} x_i y_i \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \cdot \left(\sum_{i=1}^{n} |x_i|^q\right)^{\frac{1}{q}}$$

where $\frac{1}{p} + \frac{1}{q} = 1$

$\Rightarrow$ Applying this inequality to I

$$\Leftrightarrow \log\left(\sum_{n=1}^{i} u_i^{\theta} \cdot v_i^{(1-\theta)}\right) \leq \log\left[\left(\sum_{i=1}^{n} u_i^{\theta \frac{1}{\theta}}\right)^{\theta} \cdot \left(\sum_{i=1}^{n} v_i^{1-\theta \frac{1}{1-\theta}}\right)^{1-\theta}\right]$$

$$\Rightarrow \theta \log \sum_{i=1}^{n} u_i + (1-\theta) \log \sum_{i=1}^{n} v_i$$

$\theta \mathrel{\hat{=}} \frac{1}{p}$ and $1-\theta \mathrel{\hat{=}} \frac{1}{q}$

Therefore it results:

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y) \quad \text{q.e.d.}$$

2. (Jensen's inequality) Use the definition of a concave function f, to show that $f\left(\sum_{i=1}^{m} \alpha_i x_i\right) \geq \sum_{i=1}^{m} \alpha_i f(x_i)$

where $\sum_{i=1}^{m} \alpha_i = 1 \quad \alpha_i \geq 0$

$\sum_{i=1}^{m} a_i x_i = \bar{x}$. Since f is concave, the derivative f' is monotonically decreasing.

Lets consider two cases:

(1.) $x_i \leq \bar{x}$

$$\Leftrightarrow \int_{x_i}^{x} f'(t)\,dt \geq \int_{x_i}^{x} f'(\bar{x})\,dt$$

(2.) $x_i > \bar{x}$

$$\Leftrightarrow \int_{\bar{x}}^{x_i} f'(t)\,dt \leq \int_{x}^{x_i} f'(\bar{x})\,dt$$

Through the fundamental theorem of calculus we have:

$$\int_{x_i}^{x} f'(t)\,dt = f(\bar{x}) - f(x_i)$$

$\Rightarrow$ each of the last two inequalities implies the same result:

$$f(\bar{x}) - f(x_i) \geq f'(\bar{x}) \cdot (\bar{x} - x_i)$$

this is true for all x_i, therefore:

$$f(\bar{x}) - f(x_i) \geq f'(\bar{x}) \cdot (\bar{x} - x_i)$$

$\Rightarrow a_i f(\bar{x}) - a_i f(x_i) \geq f'(\bar{x})(a_i \bar{x} - a_i x_i)$ as $a_i \geq 0$

$\Rightarrow f(\bar{x}) - \sum_{i=1}^{m} a_i f(x_i) \geq f'(\bar{x})\left(\bar{x} - \sum_{i=1}^{n} a_i x_i\right)$ as $\sum_{i=1}^{n} a_i = 1$

$\Rightarrow f(\bar{x}) \geq \sum_{i=1}^{m} a_i f(x_i)$ q.e.d. with $\bar{x} = \sum_{i=1}^{m} a_i x_i$

3. Find the minimizer of:

$$\min_{x>0,\ \boxed{\sum_{i=1}^{n} x_i = 1}} d^T x + \sum_{i=1}^{n} x_i \log x_i \qquad \text{for some given } d$$

↳ set up the Langrangian

$$L(x,\lambda) = \sum (d_i x_i + x_i \log x_i) - \cancel{\lambda}\,\mu\,(\sum x_i - 1) \qquad \text{subject to } x \geq 0$$

To minimize

↳ Stationary: $\nabla_x L(x,\lambda) = 0$ with $\log x_i = \lambda - 1 - d_i = 0$

↳ Feasability: $x > 0$ and $\sum_{i=1}^{n} x_i = 1$

It results:

$$\mu^* = -\log\left(\sum e^{(-1-d_i)}\right)$$

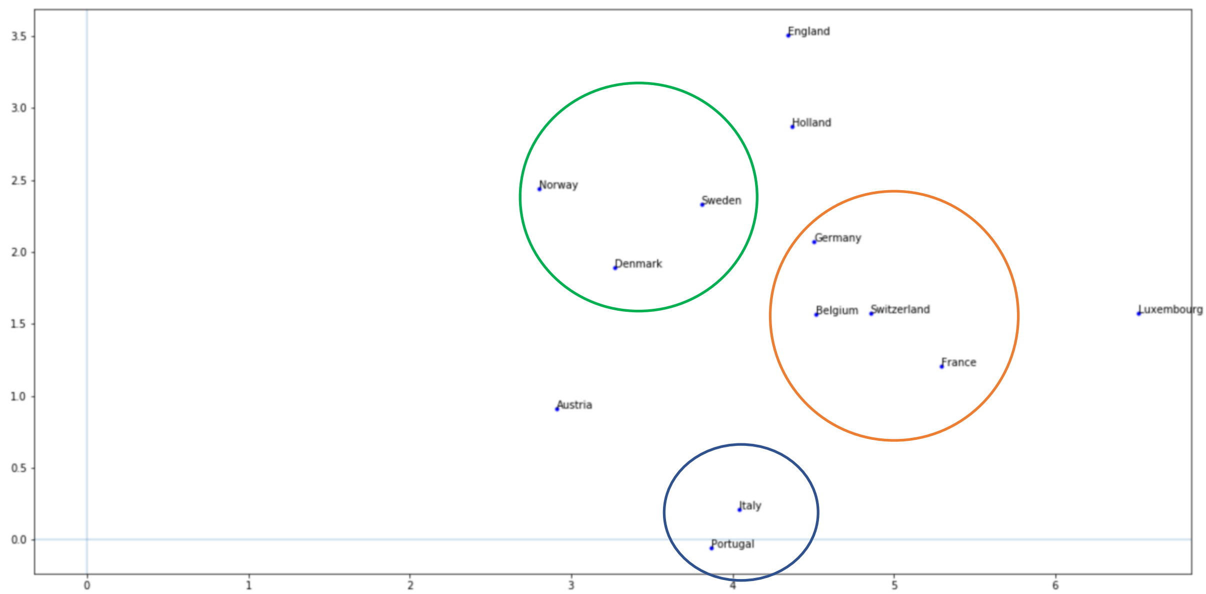
$$x_i^* = e^{(-\log(\sum_{j=1}^{n} e^{(-1-d_j)}) - 1 - d_i)}$$

Lars-Manuel Schneider 9/28/2021

ISYE-6740 Homework 2

Question 2

2. Do you observe some pattern?



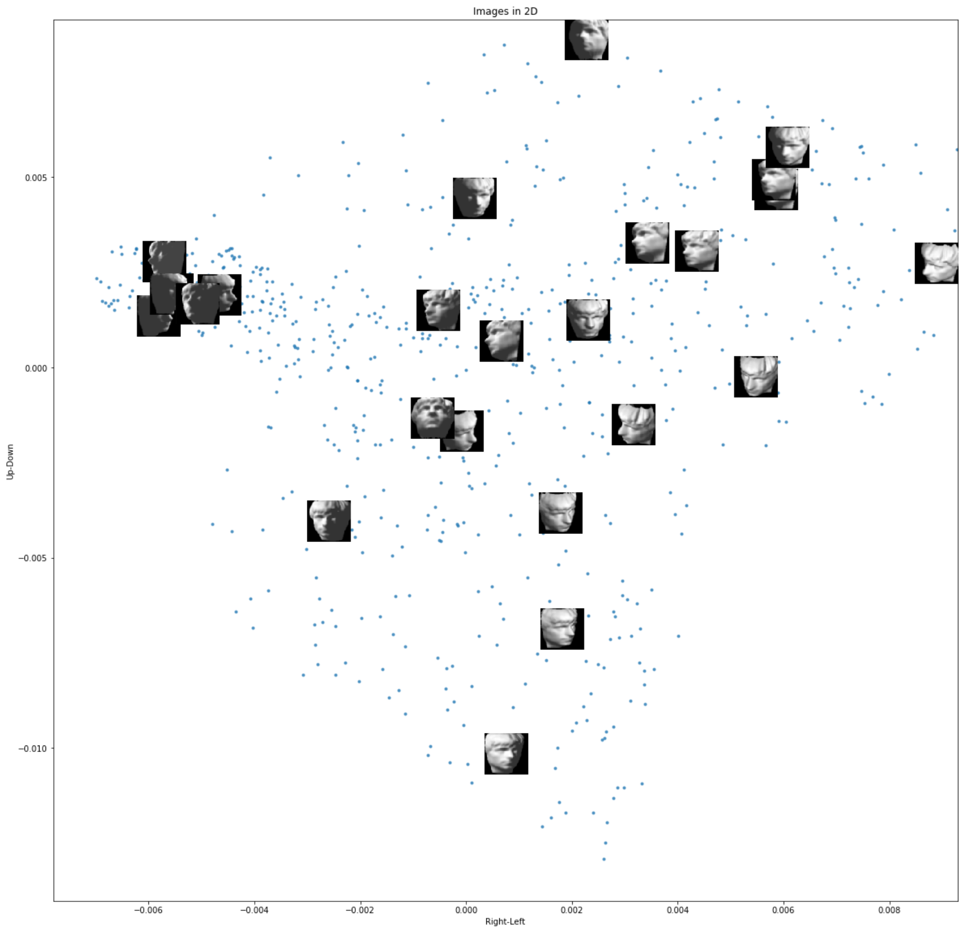
There can certainly be seen different patterns in the data. Cluster wise there can be seen, that southern countries like Italy and Portugal do have similar food cultures which makes sense. Also there are obvious similarities between Scandinavian countries like Sweden or Norway. Further, also central European countries seem to have a similar food culture. However, there are countries like Luxembourg or Austria that do not fit any cluster or other country.

Question 3

a) Show the Adjacency Matrix

```
array([[ 0.        , 18.83094952,  6.74323967, ..., 21.51126745,
         22.79289298, 18.03618033],
        [18.83094952,  0.        , 19.55307161, ..., 15.07435566,
         21.63387369, 20.97399746],
        [ 6.74323967, 19.55307161,  0.        , ..., 22.82140093,
         26.95043353, 18.68417978],
        ...,
        [21.51126745, 15.07435566, 22.82140093, ...,  0.        ,
         27.23986595, 17.19515048],
        [22.79289298, 21.63387369, 26.95043353, ..., 27.23986595,
          0.        , 20.31353772],
        [18.03618033, 20.97399746, 18.68417978, ..., 17.19515048,
         20.31353772,  0.        ]])
```

b) Implement ISOMAP with Euclidean Distance

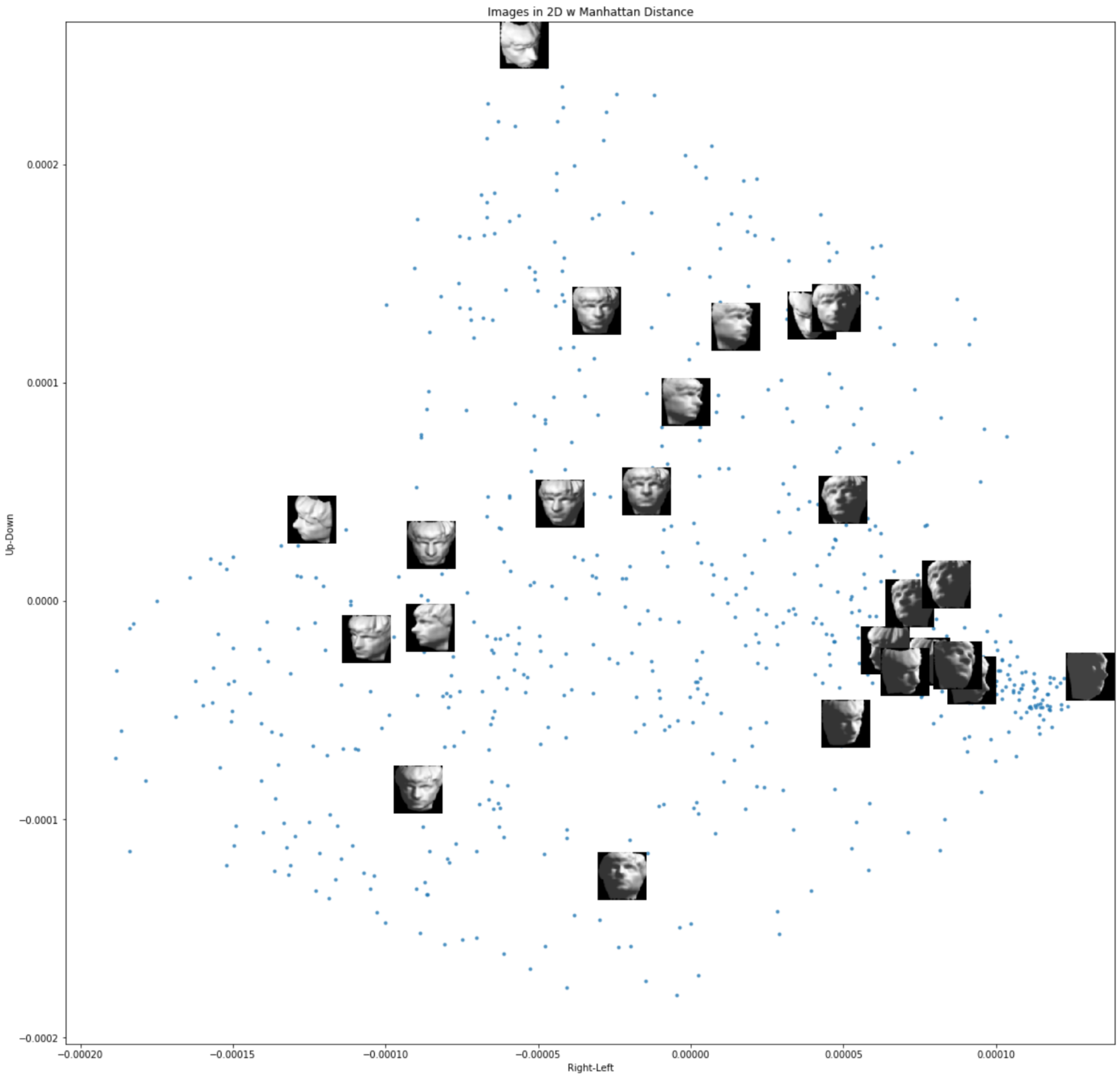


There are similarities to the arrangement in the paper. The ISOMAP algorithm generally did a good job sorting the images with the left-right and up-down pose. Particularly, the algorithm recognized the shadows on the faces decently. Occasionally the right-left pose got slightly messed up but never too bad.

Overall, ISOMAP managed to categorize the data, with the different pictures showing similar poses when close to each other.

c) Implement ISOMAP with Manhattan Distance

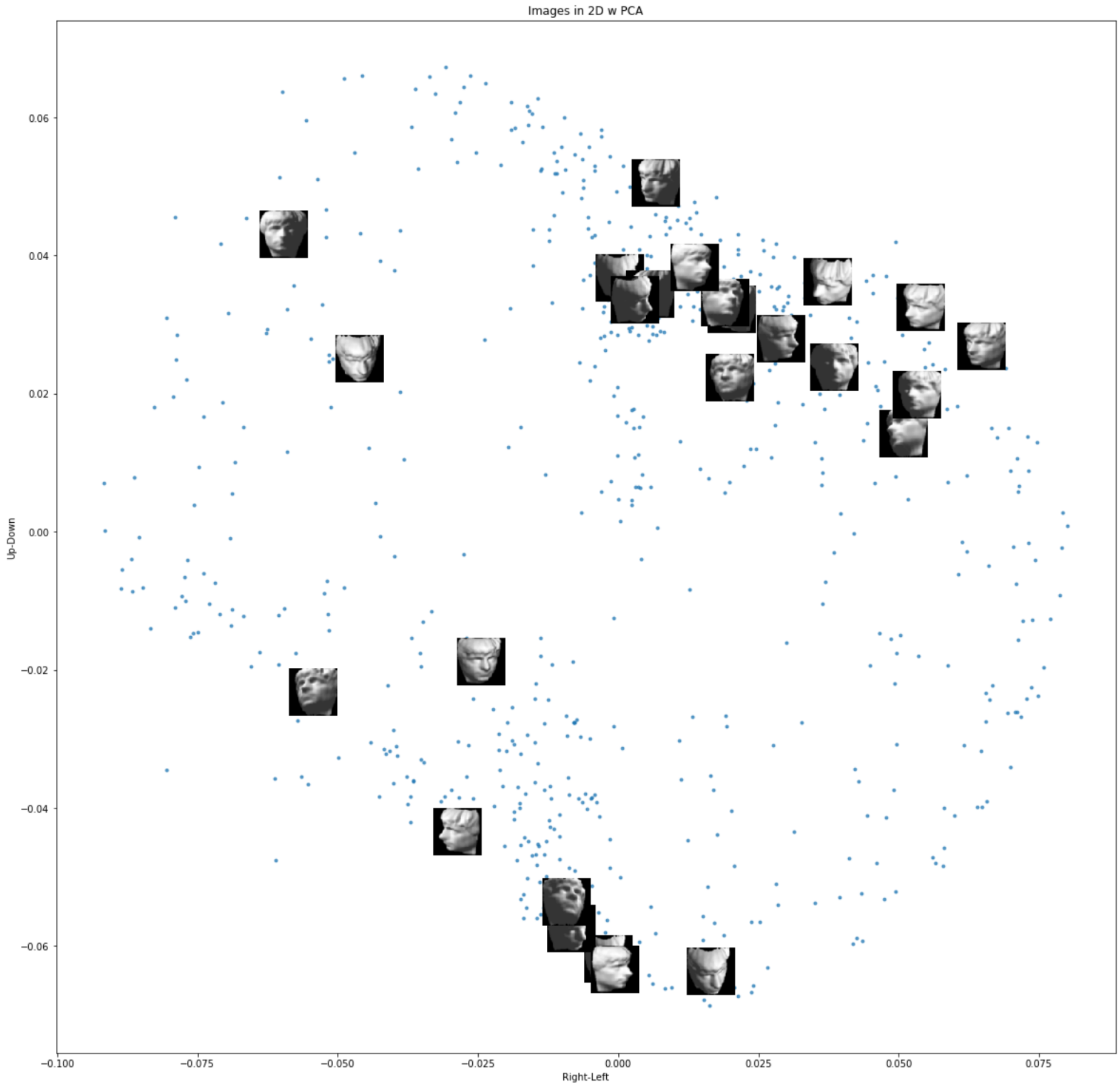
Implementing ISOMAP with Manhattan Distance definitely leads to similar results compared with when taking the Euclidean Distance. Just like in b) the algorithm categorizes the shadows on the faces accurately. The left-right poses seem slightly more confused in comparison to the Euclidean distance. Overall, the different distance metric does not lead to the same results but to similar ones.



d) Perform PCA on the images

Using PCA to categorize the images is also a proper method to do so. However, the categorization leads to comparable more errors in sorting the images to similar ones. Hence, pictures where the statue is looking left are really close to pictures where the statue looks right.

The result that ISOMAP will perform better than PCA was expected since ISOMAP is supposed to discover the nonlinear degrees of freedom that underlie complex natural observations like the images of the face in this case.


Images in 2D w PCA
Up-Down
Right-Left
0.06
0.04
0.02
0.00
−0.02
−0.04
−0.06
−0.100
−0.075
−0.050
−0.025
0.000
0.025
0.050
0.075