CMSE 820 HW 6

Hao Lin

October 25, 2018

1 Question 2

1.1 Distance by air

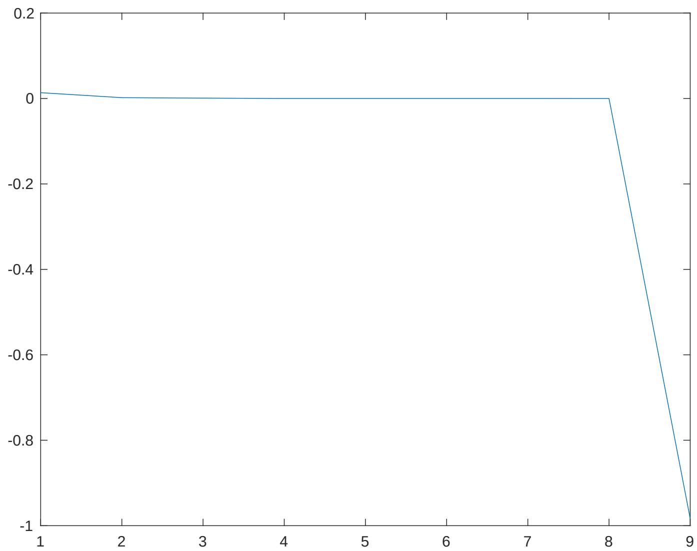

Figure 1: Normalized eigenvalues in descending order

As can be seen from the plot, not all eigenvalues are nonnegative. This is due to the fact that the distance matrix is not conditionally negative semidefinite. The distance by air is not the Euclidean distance, so the application of classical MDS may result in negative eigenvalues. For our purpose, we need only take the first two eigenvalues and eigenvectors, so the negative eigenvalues do not cause a problem.

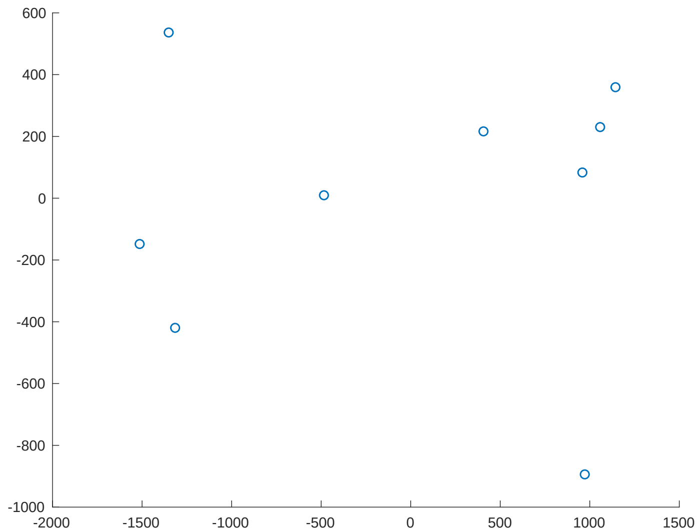

Figure 2: Scatter plot of cities using the top 2 eigenvectors.

The relative positions and distances of the cities are well-recovered.

1.2 Distance by car

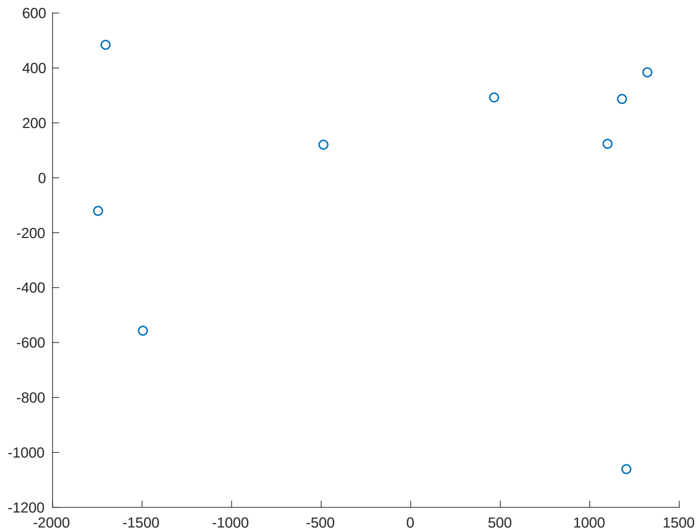

Figure 3: Scatter plot of cities using the top 2 eigenvectors.

The relative positions and distances of the cities are well-recovered.

1.3 Driving hours

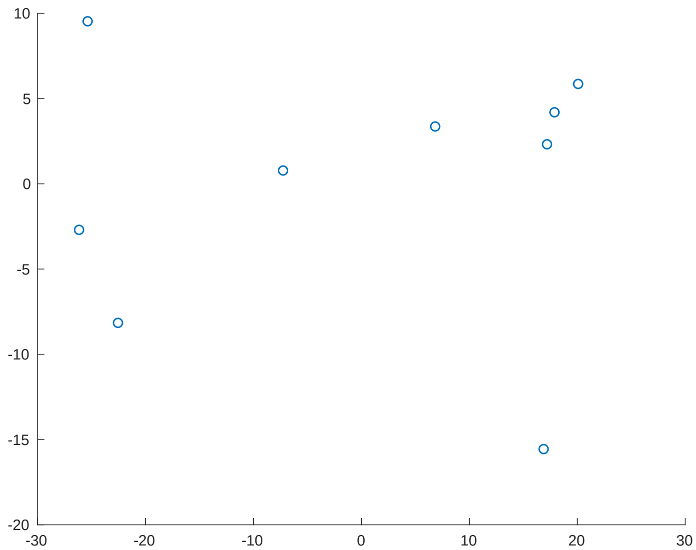

Figure 4: Scatter plot of cities using the top 2 eigenvectors.

The relative positions and distances of the cities are well-recovered.

The three scatter plots of cities using classical MDS with different kinds of distances are very similar to one another, and the results agree very well with the actual geographical distances and positions. This is because all the distances used here are very close to the Euclidean distance on a 2D map.

2 Question 3

2.1 (1)

Show that $K \succeq 0$ if and only if its eigenvalues are all nonnegative.
Proof: Since K is symmetric and hence normal, there exists an orthogonal matrix Q and a diagonal matrix Λ such that $K = Q\Lambda Q^T$.

$$\begin{aligned} & K \succeq 0 \\ \iff \quad & \forall v \in \mathbb{R}^n, v^T K v = v^T Q \Lambda Q^T v = (Q^T v)^T \Lambda (Q^T v) \geq 0 \\ \iff \quad & \Lambda \succeq 0 \\ \iff \quad & \text{All eigenvalues of } K \text{ are nonnegative.} \end{aligned}$$

2.2 (2)

Show that $d_{ij} = K_{ii} + K_{jj} - 2K_{ij}$ is a squared distance function, i.e., there exists vectors $u_i, u_j \in \mathbb{R}^n$ for $1 \leq i, j \leq n$ such that $d_{ij} = \|u_i - u_j\|^2$.
Proof: Let $K = Q\Lambda Q^T$ be the eigenvalue decomposition of K. Define $U = \Lambda^{\frac{1}{2}} Q^T$. In particular, $K = U^T U$, i.e., $\forall i, j,\ K_{ij} = u_i^T u_j$ where u_i and u_j are the i-th and the j-th columns of U, respectively.

$$d_{ij} = u_i^T u_i + u_j^T u_j - 2u_i^T u_j = (u_i - u_j)^T (u_i - u_j) = \|u_i - u_j\|^2.$$

Thus, d is a square distance function.

2.3 (3)

Show that $A + B \succeq 0$ and $A \circ B \succeq 0$ (Hadamard product). Show that the eigenvalues of AB are all positive.
Proof:

$$\forall v \in \mathbb{R}^n,\ v^T (A + B) v = v^T A v + v^T B v \geq 0 \Rightarrow A + B \succeq 0.$$

Let the eigenvalue decompositions of A and B be $A = U\Lambda U^T = \sum_i \lambda_i u_i u_i^T$ and $B = V\Gamma V^T = \sum_i \gamma_i v_i v_i^T$.

$$A \circ B = \sum_{i,j} \lambda_i \gamma_j (u_i u_i^T) \circ (v_j v_j^T) = \sum_{i,j} \lambda_i \gamma_j (u_i \circ v_j)(u_i \circ v_j)^T.$$

Note that $\forall x \in \mathbb{R}^T, x^T (u_i \circ v_j)(u_i \circ v_j)^T x = ((u_i \circ v_j)^T x)^T ((u_i \circ v_j)^T x) \geq 0$, so $\forall i$, j, $(u_i \circ v_j)(u_i \circ v_j)^T \succeq 0$. It follows that the nonnegative sum $A \circ B$ of positive semi-definte matrices is also positive semi-definite.

2.4 (4)

Proof: Let v be an arbitrary eigenvector of AB corresponding to the eigenvalue λ, i.e., $ABv = \lambda v$. Assume $A \succ 0$ and $B \succ 0$.

$$ABv = AB^{\frac{1}{2}} B^{\frac{1}{2}} v = \lambda B^{-\frac{1}{2}} B^{\frac{1}{2}} v \implies (B^{\frac{1}{2}} A B^{\frac{1}{2}})(B^{\frac{1}{2}} v) = \lambda (B^{\frac{1}{2}} v).$$

That is, an eigenvalue of AB must be an eigenvalue of $B^{\frac{1}{2}} A B^{\frac{1}{2}}$. Note that $\forall x \in \mathbb{R}^n$, $x^T B^{\frac{1}{2}} A B^{\frac{1}{2}} x = (B^{\frac{1}{2}} x)^T A (B^{\frac{1}{2}} x) > 0$. So all eigenvalues of AB are positive.

2.5 (5)

If $A \succeq 0$ and $c \geq 0$, then $cA \succeq 0$.
Proof:

$$\forall x \in \mathbb{R}^n, x^T c A x = c x^T A x \geq 0 \implies cA \succeq 0.$$

2.6 (6)

If $A \succeq 0$ and C can be writeen as $C = \{t_{[i]}A_{ij}t_{[j]}\}$, then $C \succeq 0$.
Proof:

$$\forall x \in \mathbb{R}^n, \sum_{i,j} x_i x_j C_{ij} = \sum_{i,j} (x_i t_{[i]}) A_{ij} (x_j t_{[j]}) \geq 0 \implies C \succeq 0.$$

2.7 (7)

Show that the Hadamard integral power $A^{\circ p} = \{A_{ij}^p\}$ with $p \in \mathbb{N}$ and the Hadamard exponential $\exp(\circ A) = \{\exp(A_{ij})\}$ are p.s.d.
Proof: By (3), $A \succeq 0 \implies A^{\circ 2} = A \circ A \succeq 0$. By induction, $\forall p \in \mathbb{N}, A^{\circ p} \succeq 0$.

$\exp(\circ A) = \sum_{p=0}^{\infty} \frac{A^{\circ p}}{p!} \succeq 0$. Note that the finite sum $S_n(A) = \sum_{p=0}^{n} \frac{A^{\circ p}}{p!}$ is p.s.d, and $\{S_n\}$ converges uniformly to $\exp(\circ A)$.

3 Question 4

3.1 1. Infinitely Divisible Kernels

Let C be a symmetric matrix, and define $B = \exp(\circ - C)$. Show that B is infinitely divisible if and only if C is conditionally negative definite.
Proof: $(\Rightarrow)$ $\forall \lambda > 0$, Define $A(\lambda) = \frac{1}{\lambda}(\mathbf{1}\mathbf{1}^T - B^{\circ \lambda})$. Note that $\lim_{\lambda \to 0} A(\lambda) = C$.

$$\forall x \text{ such that } \mathbf{1}^T x = 0,\ x^T A x = \frac{1}{\lambda}(x^T \mathbf{1}\mathbf{1}^T x - x^T B^{\circ\lambda} x) = -\frac{1}{\lambda} x^T B^{\circ\lambda} x \leq 0 \implies A(\lambda) \text{ is c.n.d.}$$

It follows that C is c.n.d.

$(\Leftarrow)$ Define $F = -H_\alpha C H_\alpha^T$. By construction, F is p.s.d. regardless of the choice of H_α. In particular, let's consider the case where $H_\alpha = I - \frac{\mathbf{1}\mathbf{1}^T}{n}$.

$$F_{ij} = -\left(C_{ij} - \frac{1}{n}\sum_j C_{ij} - \frac{1}{n}\sum_i C_{ij} + \frac{1}{n^2}\sum_{ij} C_{ij} \right).$$

Note that $\exp(\circ F)$ is p.s.d, and

$$[\exp(\circ F)]_{ij} = \exp(F_{ij}) = \exp(-S)\exp(\bar{C}_{i,})\exp(-C_{ij})\exp(\bar{C}_{.,j}),$$

where $S = \frac{1}{n^2}\sum_{ij} C_{ij}$, $\bar{C}_{i,.} = \frac{1}{n}\sum_j C_{ij}$ and $\bar{C}_{.,j} = \frac{1}{n}\sum_i C_{ij}$.

Define $M = \operatorname{diag}\{\exp(-\bar{C}_{i,.})\} \succeq 0$ and $N = \operatorname{diag}\{\exp(-\bar{C}_{.,j})\} \succeq 0$ and rewrite $B = \exp(\circ - C)$ as

$$B = \frac{1}{\exp(-S)} M \exp(\circ F) N \succeq 0.$$

The infinitely divisibility of B can be seen by

$$\forall r > 0,\ B^{\circ r} = \frac{1}{\exp(-rS)} M^r [\exp(\circ F)]^{\circ r} N^r = \frac{1}{\exp(-rS)} M^r \exp(\circ r F) N^r \succeq 0.$$

3.2 2 Gaussian Kernel

$D^{(2)}$ is a sqaured Euclidean distance matrix implies $\exists x_i \in \mathbb{R}^p$ for $i = 1, ..., n$ such that $D^{(2)}_{ij} = x_i^T x_i - 2x_i^T x_j + x_j^T x_j$. $\forall \lambda \geq 0$,

$$\exp(-\lambda D^{(2)}_{ij}) = \exp(-\lambda x_i^T x_i) \exp(2\lambda x_i^T x_j) \exp(-\lambda x_j^T x_j)$$

$$\implies \exp(\circ - \lambda D^{(2)}) = \Lambda \exp(\circ 2\lambda X^T X)\Lambda \succeq 0,$$

where $\Lambda = \mathrm{diag}\{-\lambda x_i^T x_i\}$ and $X^T X \succeq 0$.

To show that $\tilde{D}^{(2)}_\lambda$ is a squared distance matrix, it suffices to show that $\tilde{D}^{(2)}_\lambda$ is c.n.d. with all diagonal entries equal to 0.

$$\forall\, i,\ \tilde{D}^{(2)}_{\lambda,ii} = 1 - \exp(-\lambda D^{(2)}_{ii}) = 1 - \exp(0) = 0.$$

We can rewrite $\tilde{D}^{(2)}_\lambda$ as

$$\tilde{D}^{(2)}_\lambda = \mathbf{1}\mathbf{1}^T - \exp(\circ - \lambda D^{(2)}).$$

$$\forall\, x \text{ such that } \mathbf{1}^T x = 0,\ x^T(\mathbf{1}\mathbf{1}^T - \exp(\circ - \lambda D^{(2)})x = -x^T \exp(\circ - \lambda D^{(2)})x \leq 0,$$

because $\exp(\circ - \lambda D^{(2)})$ is p.s.d. Thus, $\tilde{D}^{(2)}_\lambda$ is c.n.d. (with a zero diagonal).