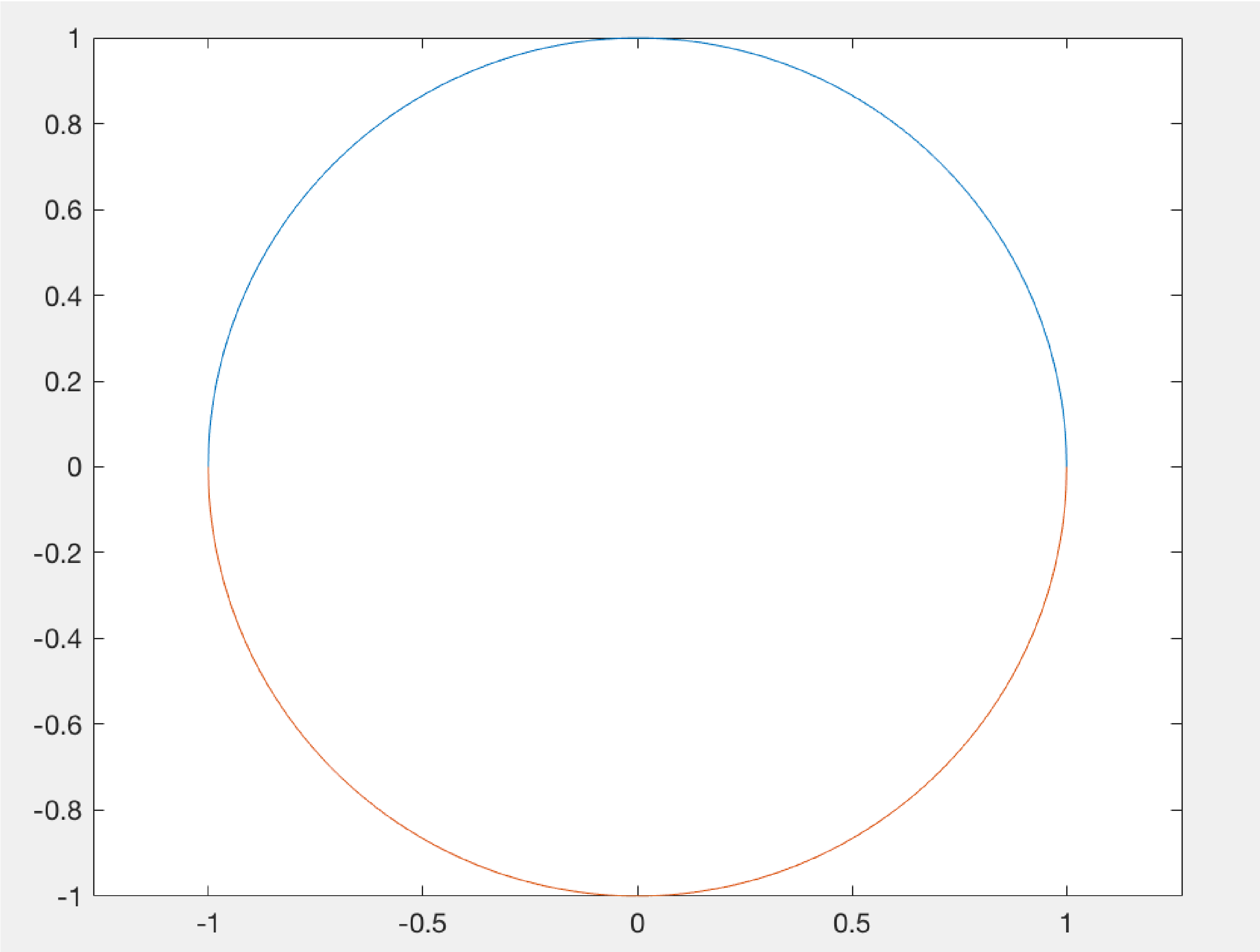
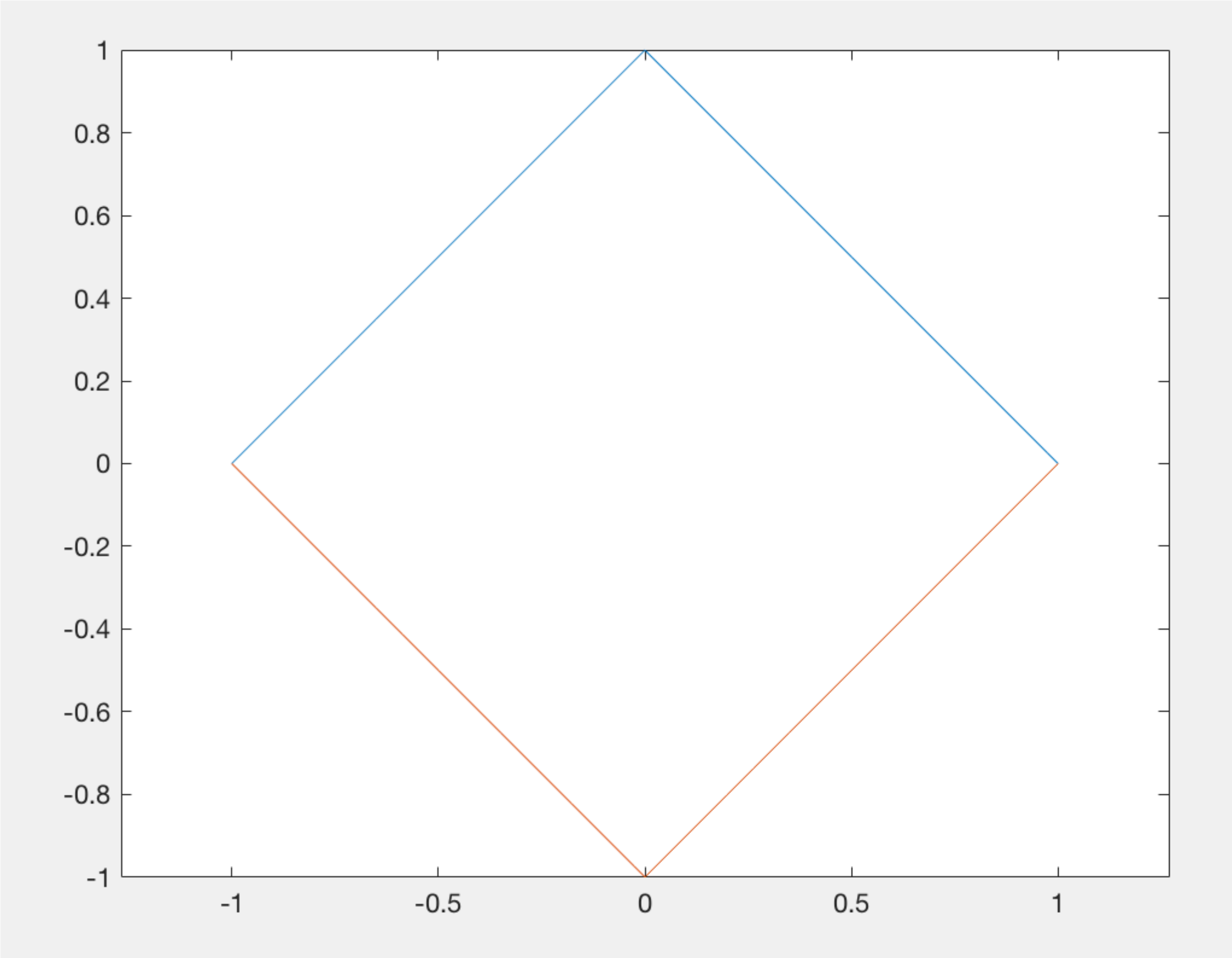
DM - Assignment 02

Problem 1

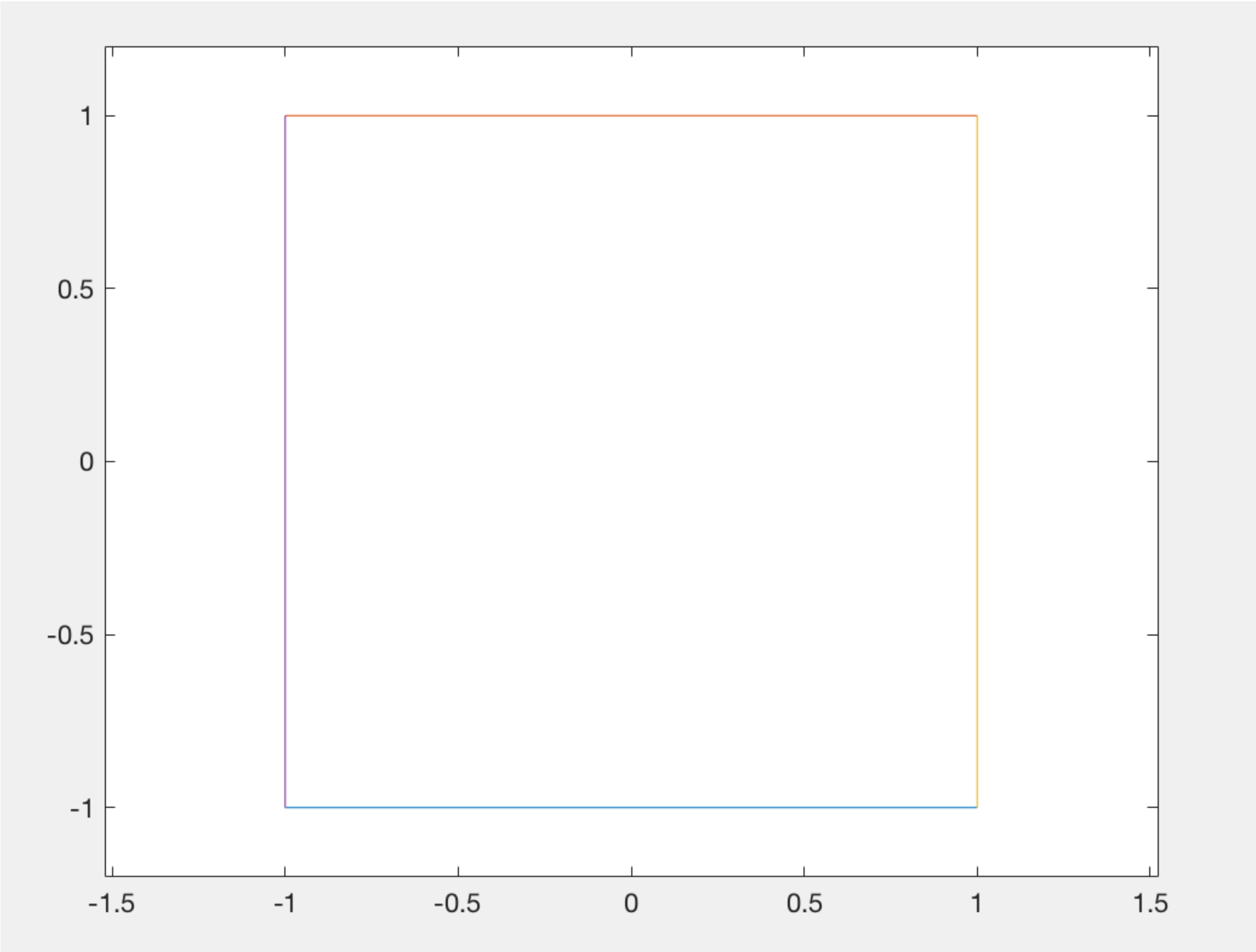
1. Euclidean



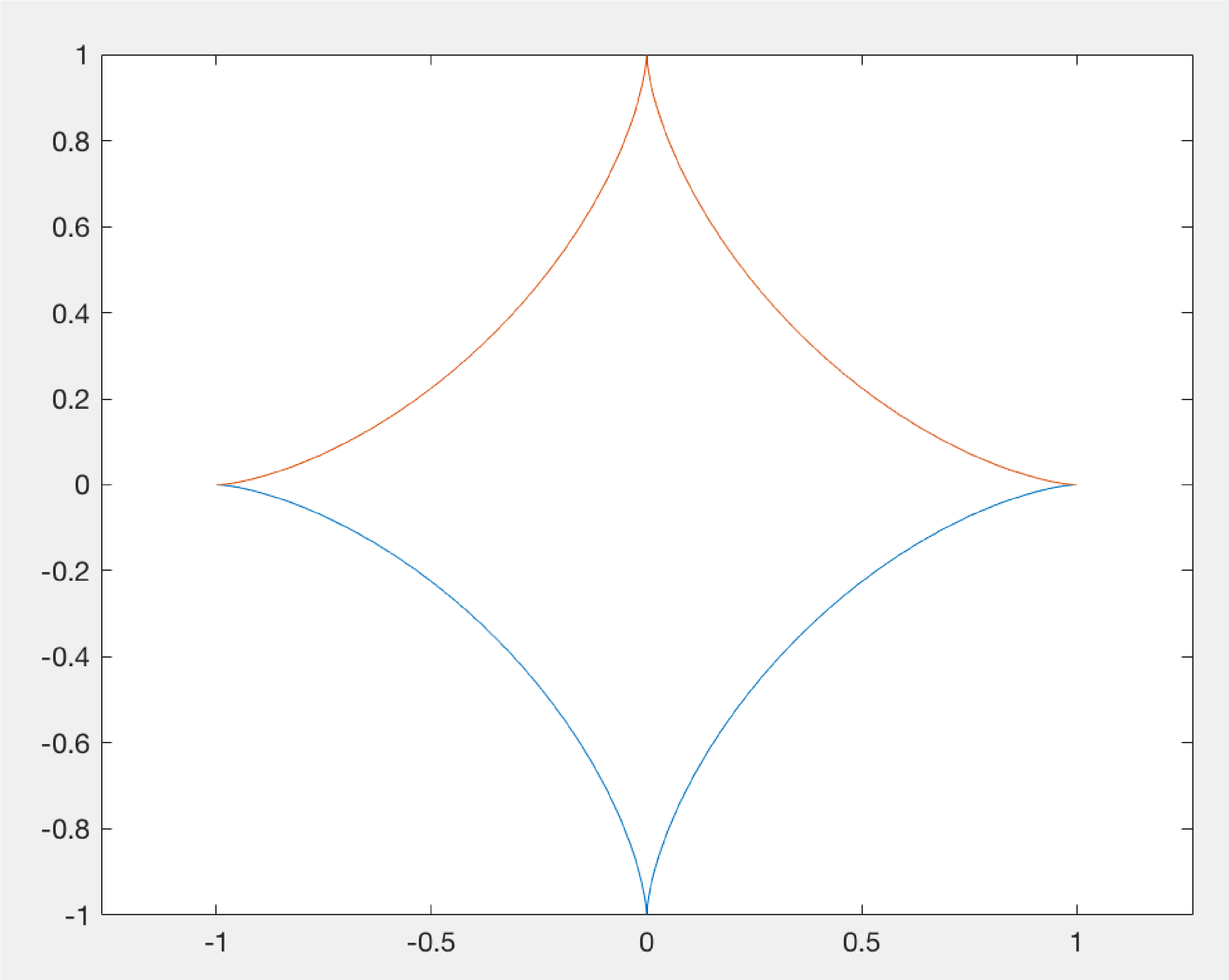
2. Manhattan



3. Supremum

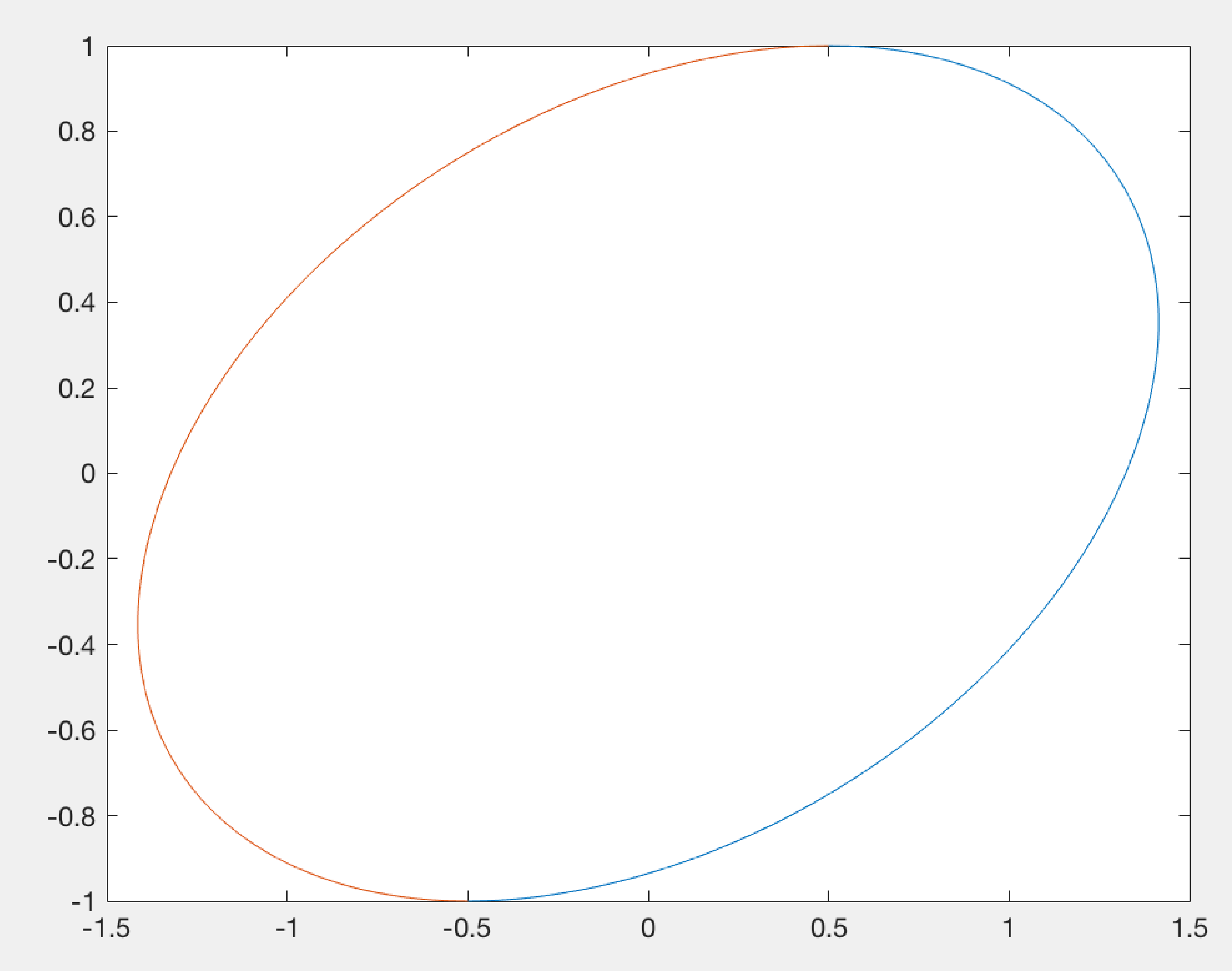


4. $L^{2/3}$



5. Mahalanobis

Represent x and y in a row vector and obtain `x^2-xy+2y^2=7/4`, `x = 0.5y +- (7/4 * (1-y^2))^0.5`.



Problem 2

1. 1 - J(x,y)

- If d(x,y) = 1 - J(x,y) = 0, then J(x,y) = 1, |A∩B| = |A∪B|, A must be the same set as B; and the reversed deduction from A=B to |A∩B| = |A∪B| -> d(x,y) = 0 is quite straight forward.
- Because J(x,y) is symmetric to x and y (say A and B), so d(x,y) is also symmetric.
- Randomly pick some set z in the definition space, we need to prove d(x,y) ≤ d(x,z) + d(z,y), which is also interpretated as J(x,z) + J(z,y) - J(x,y) ≤ 1; let's denote the whole area of set x,y,z as X,Y,Z, respectively, and the cross part of XY,YZ,ZX as P,Q,R, the lapped middle part area of all three sets denoted also as W; so the formula above could be transformed into R/(Z+X-R) + Q/(Y+Z-Q) - P/(X+Y-P) ≤ 1,

2. arccos((x•y) / |x||y|)

- If d(x,y) = arccos((x•y) / |x||y|) = 0, then (x•y)/|x||y| = 1 as the domain of arccos function is `[-1,1]`, meaning the angle between x,y is zero, they are in the exact same direction; recalling they are both vectors, so x and y are equivalent; vice versa, if x=y, then the angle between is certainly 0 and cosine similarity is 1, thus the fact aht d(x,y) = 0 also stands.

- Transforming d(x,y) into form of arccos((x•y) / |x||y|) clear indicates its symmetry between x and y, meaning d(x,y) = d(y,x).

- Randomly pick some set z also as a vector in x and y's space, we need to prove d(x,y) ≤ d(x,z) + d(z,y), where d(•,•) belongs to `[0,π]`. If z does not belong to span{x,y}, project it to the plane which x and y spans, and d(z,x) or d(z,y) will always decrease (could be proved using cosine theorem). Now we only need to prove the case that z is within span{x,y}. Notice that if z lies between the direction of x and y, then d(x,y) = d(x,z) + d(z,y), the statement stands; otherwise if z lies outside to the minimal angle which formed by x and y, say if direction of z is closer to y without loss of generality, then d(x,z) > d(x,y) simply proved true, and the original inequality also stands; till now we have already covered all situations that could exist in this equality.

3. Length of Shortest Path in G(V,E,w)

- If d(x,y) = 0 then the sum of weight coefficients along the short path between x and y is 0, but regarding each edge has a positive value of weight, meaningful no edge is actually walked through, x=y; vice versa, x=y can be easily reducted to d(x,y) = 0.

- Since the graph that this kind of metric based is undirected, so shortest path from x to y should be exactly the same shortest path from y to x, making it symmetrical and thus d(x,y) = d(y,x).

- For any z in G(V,E,w), because d(x,y) is already defined as the shortest path sum from x to y, so the path sum from x->z->y, which is d(x,z) + d(z,y) could not be less than that, meaning d(x,y) ≤ d(x,z) + d(z,y); the equality condition only stands when z is on the shortest path between x and y.

4. Mutual Set Difference

- If d(A,B) = 0 then d(A/B) = d(B/A) = 0 cuz they are both non-negative, meaning A is equivalent to B; reversely, if A = B, d(A,B) = d(A/B) + d(B/A) = 0 + 0 = 0.

- d(A,B) = d(A/B) + d(B/A) = d(B/A) + d(A/B) = d(B,A).

- For arbitrary set C, let's denote the cross part of AB,BC,CA as P,Q,R, respectively, and the lapped middle part area of all three sets denoted also as W; now d(B,C) + d(C,A) - d(A,B) = 2(C+P-Q-R) = 2((P-W) + (C-A-B)); because norms of both sets, namely P-W and C-A-B cannot be negative, so the formula above has a value that is non-negative, thus the triangle rule of distancr metric is proved.

Problem 3

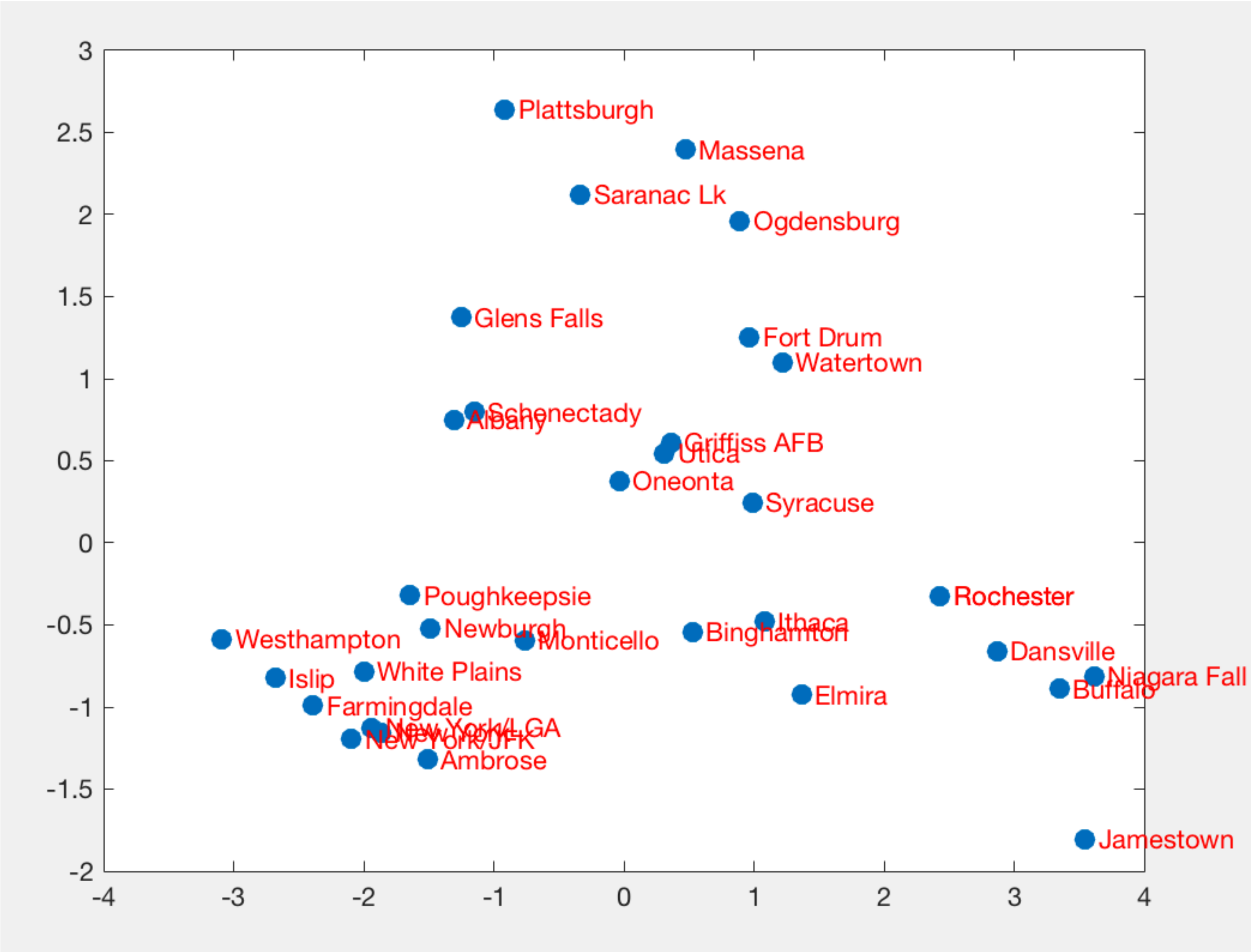
As our data samples, X which has m rows indicating number of samples while n colums indicating number of variable or features is mean substracted, meaning X = JX already stands, define J = I - (1/m) * ones(3,3).

In PCA, we first obtain covariance matrix by Cov = (1/m-1) * X'X, also Cov = USU' after SVD (since Cov is symmetric), where columns of U are eigenvectors of X, diag(S) as sequence of eigenvalues.

Now move the problem to Classical MDS, if there's no distance matrix D given, we should discover the relationship between D and X. Let us denote X_i as the i-th row of X or the i-th sample, then as a pair-wise distance, $D_{i,\,j} = ||X_i - X_j||^2 = ||X_i||^2 + ||X_j||^2 - 2 <X_i, X_j>$ according to cosine theorem. From that we can generate D = Z - 2XX' + Z', where we define $Z_{i,\,j} = ||X_i||^2$ that makes Z satisfies the JZ = O. Multiply a J to left hand sides of both sides of the equation above, we eliminate Z and obtain JD = -2XX' since JX=X; one more time multiply a J' to the right hand side and we are done by obtaining Gram = -1/2 * JDJ' = XX', the exact matrix used for SVD in MDS.

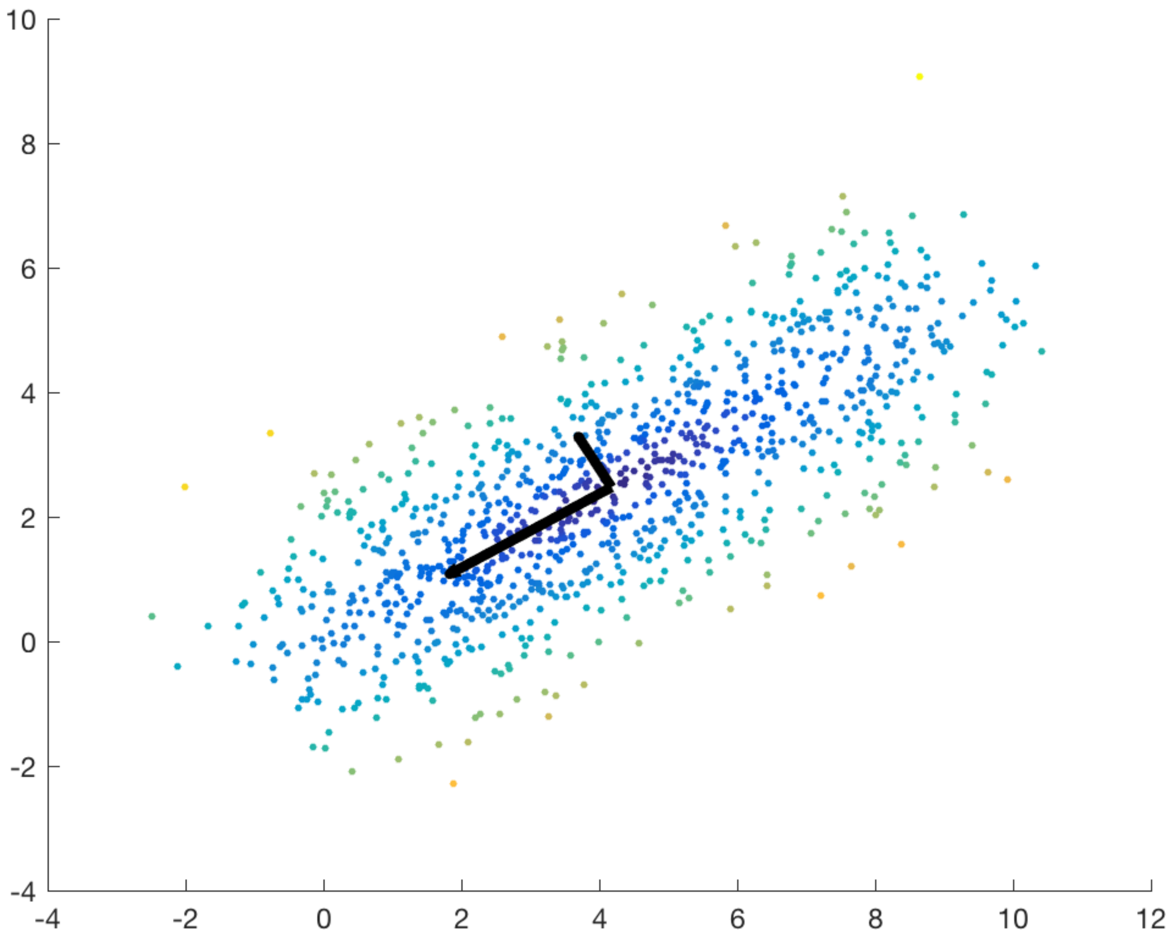
To again make it clear, in MDS, Gram = XX' = (m-1) * Cov' = (m-1) * USU' according to previous SVD. Taken Gram into its own SVD process and we obtain Gram = VßV', then U and V are only different in some scaling coefficients, while U_{reduce} = U(:, 1:k) as prime eigenvectors are just in scale to V(:, 1:k), which is the first k columns of V, and k ≤ (number of elements in diag(S) that is not zero). Thus, the projection results of PCA and Classic MDS, $X_{proj} = XU_{reduce}$ and $Y = Vß^{1/2}$ respectively, are equivalent.

Problem 4



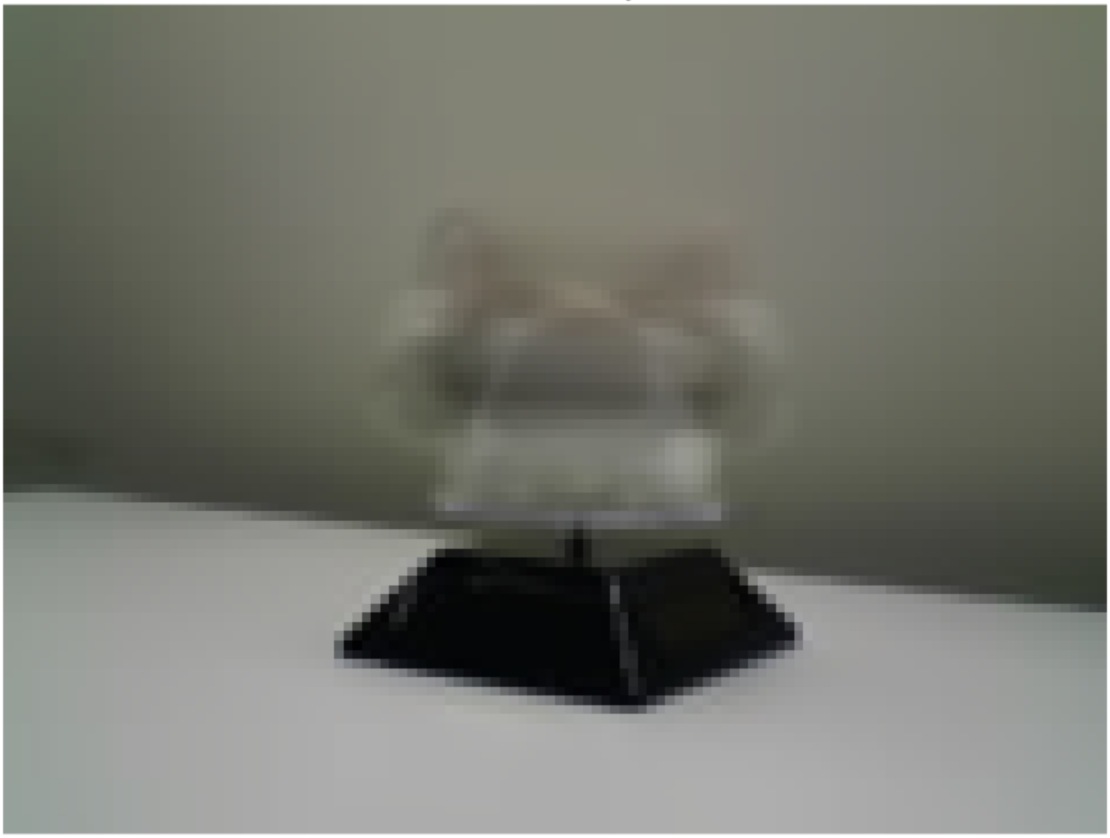
3
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-4
-3
-2
-1
0
1
2
3
4
Plattsburgh
Massena
Saranac Lk
Ogdensburg
Glens Falls
Fort Drum
Watertown
Schenectady
Albany
Griffiss AFB
Utica
Oneonta
Syracuse
Poughkeepsie
Rochester
Westhampton
Newburgh
Monticello
Binghamton
Ithaca
Dansville
Islip
White Plains
Niagara Fall
Buffalo
Elmira
Farmingdale
New York/LGA
New York/JFK
Ambrose
Jamestown

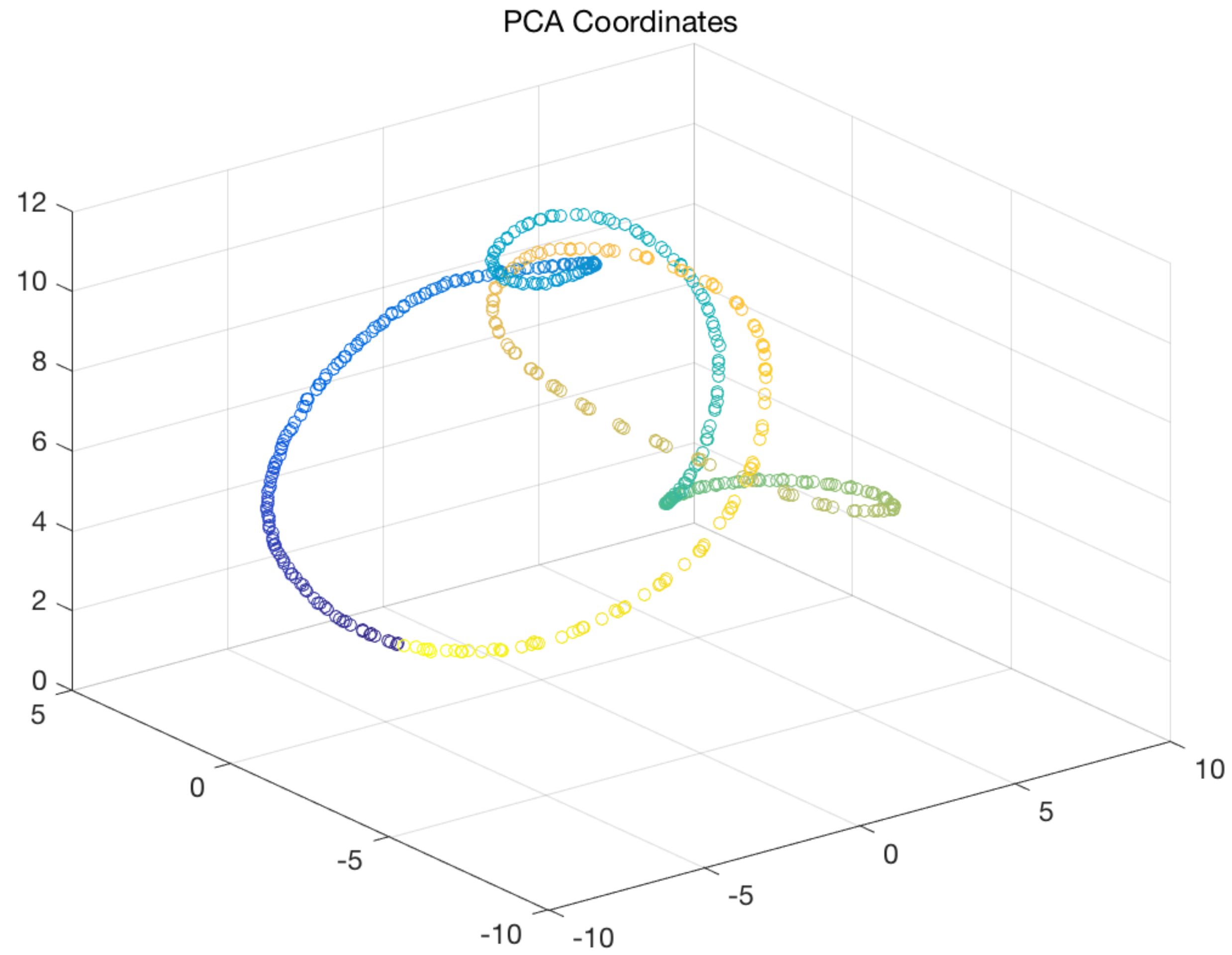
Problem 5

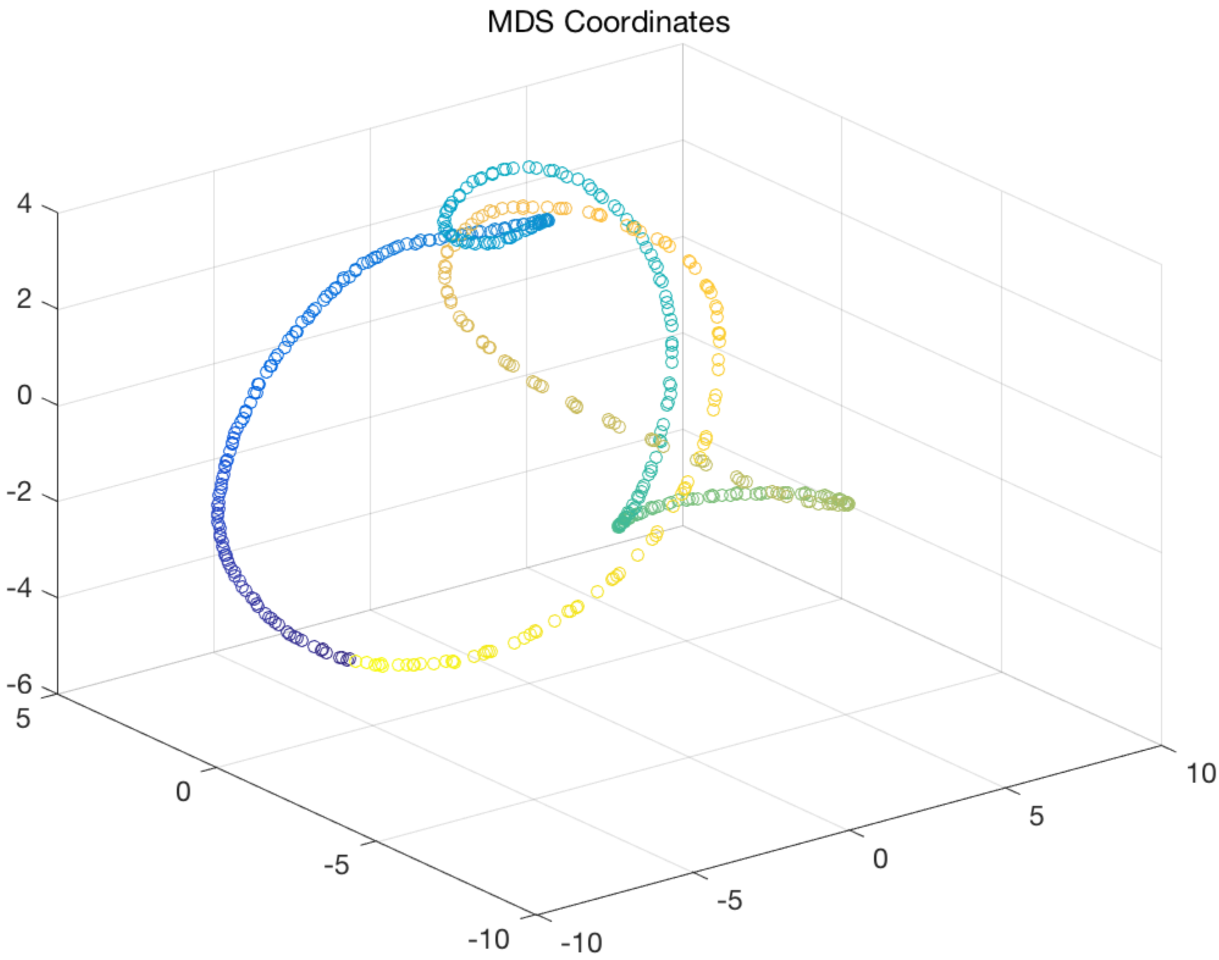


Problem 6

Mean bunny




PCA Coordinates
12
10
8
6
4
2
0
5
0
-5
-10
-10
-5
0
5
10


MDS Coordinates
4
2
0
-2
-4
-6
5
0
-5
-10
-10
-5
0
5
10