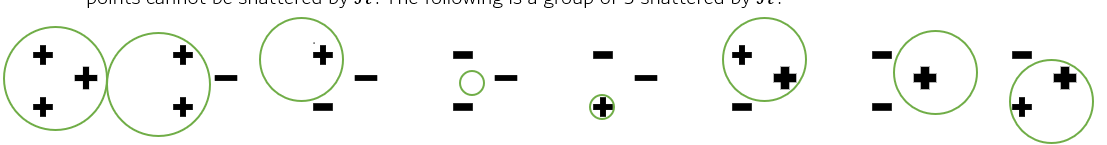
**Problem 1**

Question 1

All we need to prove is that there exists a group of 3 points shattered by , but every group of 4 points cannot be shattered by . The following is a group of 3 shattered by .



Now let's look at a set of 4 points, . There could be 2 cases.

1. 3 colinear points, therefore the middle one cannot be assigned – with the sides being labeled +
2. If the points are convex, according to theorem 1 (look at appendix 1), some non-adjacent pair of points cannot be covered (assigned +) without covering some third point (which we want to be -).
3. If there is a middle point, then assigning - to the middle one but + to the rest is impossible to be achieved.

Note: diagrams are disgusting sorry 😊

Question 2

Question 3

The problem is now PAC learnable, therefore, according to the fundamental theorem of statistical learning:

**Problem 4**

Section a

We want to find local minimum, which can be done by finding points where function gradient of is 0.

Section b

I added two lines of code which initialize a Kmeans object, and fits it to the data.

k = 2

clust = Kmeans(k)

clust.fit(data)

I also added a function that reassigns the centroids after slef.labels has been updated by the class

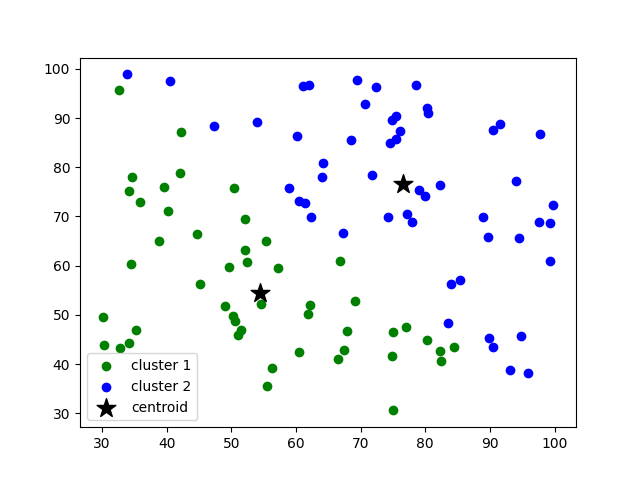
    def reassign\_centroids(self, X, labels):

        centroids = np.zeros((self.n\_clusters, X.shape[1]))

        for i in range(self.n\_clusters):

            centroids[i] = np.mean(X[labels==i])

        return centroids

here are the results:

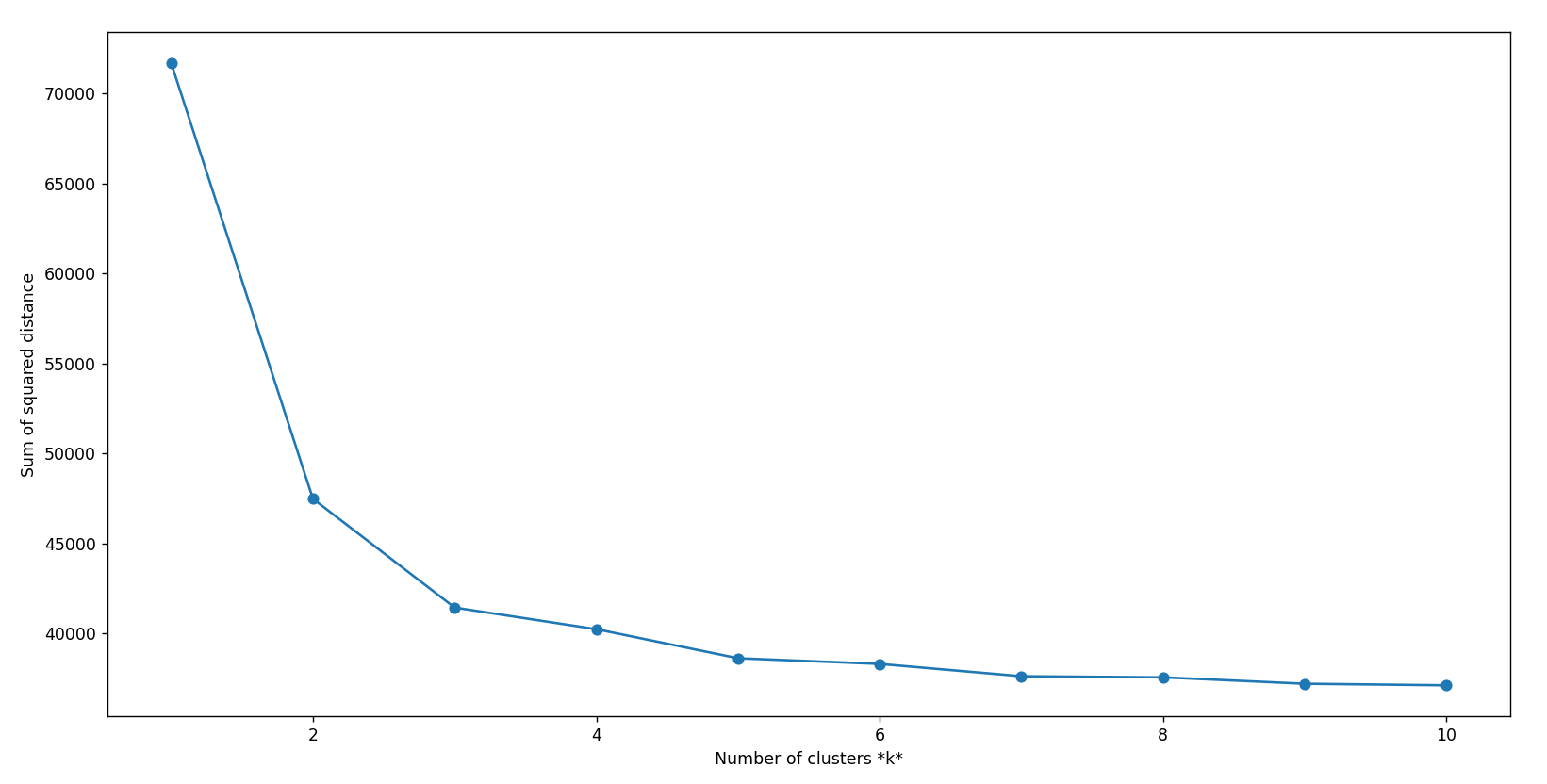
and using the elbow function we plotted the inertia of the algorithm as a function of k.

for k in list\_k:

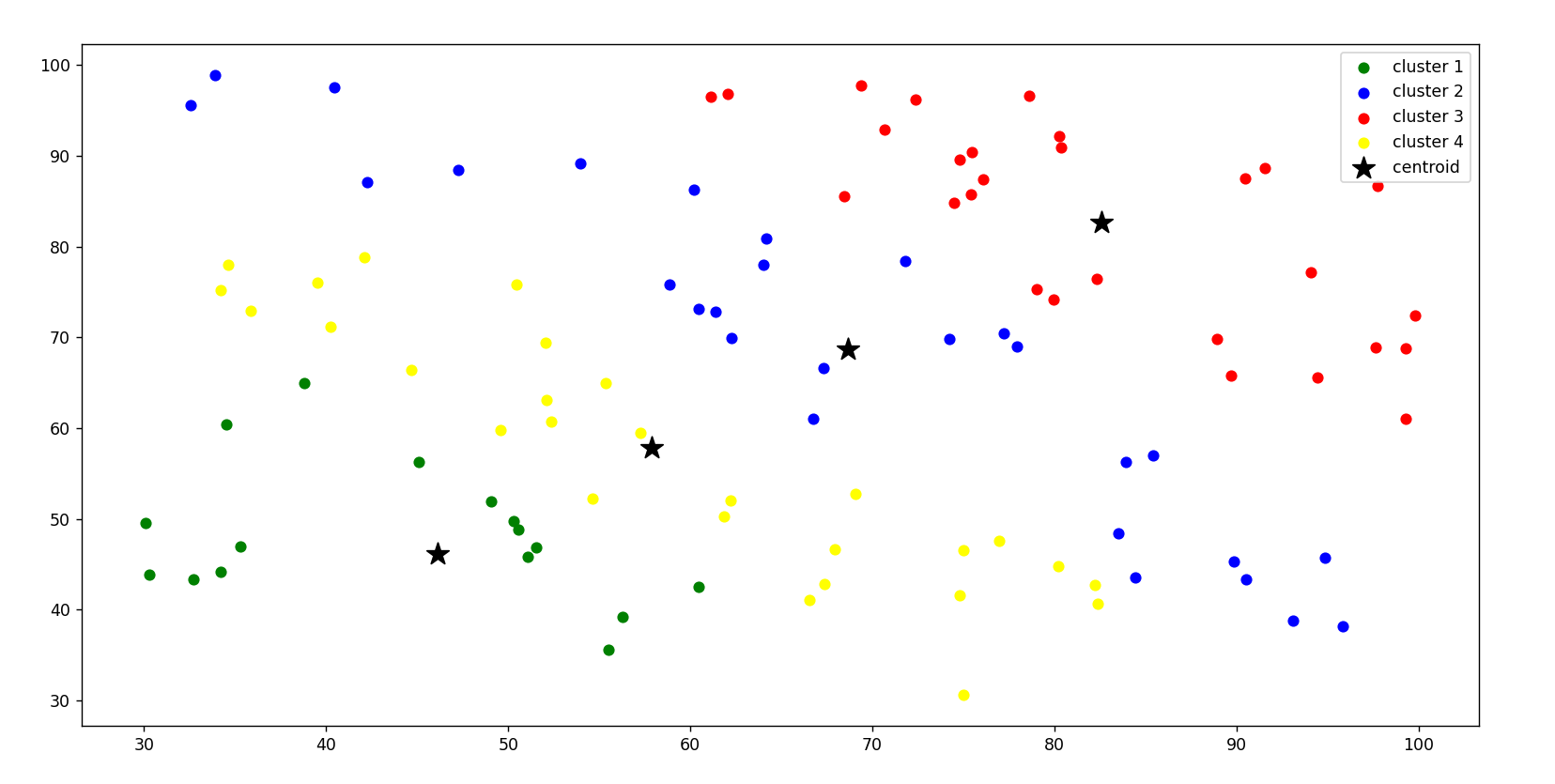
    clust = Kmeans(k)

    clust.fit(data)

    sse.append(clust.compute\_sse(data,clust.labels,clust.centroids))



And therefore the best choice is k=4 or k=3 (somewhat between overfitting and underfitting, high bias and high variance, which is right after the drastic change). Here are the results of k=4.



Section D

I used sickit-learn class for implementing k-means, and found clustering for 20 colors. Then, I built a function that takes the k-means object and the picture, and for each pixel, maps the compressed color. Then we plotted both pictures. Here is the code and the output.

# returns compresseed image from image X by cluster clust

def compress(X,clust):

    for i in range(20):

        X[clust.labels\_==i] = clust.cluster\_centers\_[i]

    return X



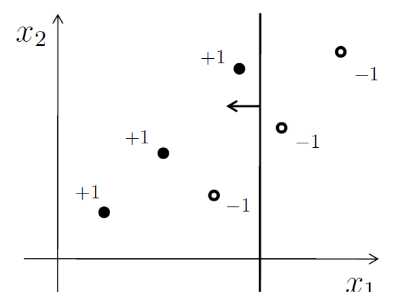
**Problem 5**

1. Initializing we get that , and
2. For misclassified points, .

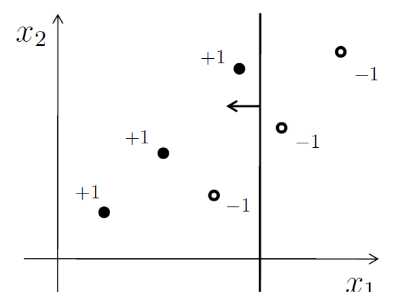
For well classified points,

To normalize we divide by , and get

1. We lose one positive point for the expensive negative point, and get that:



1. , and since is a monotonically decreasing function, .
2. We proved that if we choose the same weak classifier with , we will get . This can be proven to be correct with the blue weak classifier, in which , which is not the case for green classifier.
3. We draw all of them and take the weighted sum sign:



Train accuracy rate is 100% since all samples are eventually classified correctly.

**Appendix 1 – Theorem 1**

**Definition.1** let . An area is -dominant with respect to if

Dominance area of with respect to denoted as is defined as:

**Definition.2** let . dominates if:

**Theorem.1** Let be a convex set of points, then there's a partition of into where , and dominates .

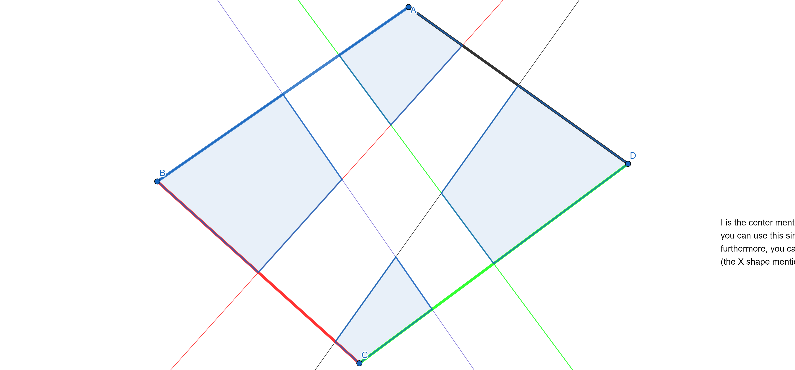
**Proof.1** first, we look at polygon (assuming without loss of generality this is their order in the space). We build perpendicular bisectors of each edge of this polygon. Notice that these bisectors are the sub-space at which distances are equal from the vertices of the edge. Therefore, they define areas at which inequalities between distances from any point inside them are ensured. More specifically, they define dominance areas of each point with respect to their adjacent vertices.

If the points are convex, we get that the quadrilateral is built in which the perpendicular bisectors of adjacent edges to the same point form adjacent edges in the new quadrilateral. This means that according to lemma 1 two facing points (w.l.o.g.) will apply:

And therefore, their union applies:

Note this proves that assigning + to and to will never be satisfied for any center .

**Lemma.1** let be a convex set of points then two non-adjacent points will have mutually exclusive dominance areas compliments.

**Proof.1** assume by contradiction all pairs don't have mutually exclusive dominance areas compliments. We know for sure that the shaded areas in the following diagram are not of the dominance area of each point.

Any extension of the compliment must be in units of bisectors, which means that in order for the compliments to have shared areas, they must include the quadrilateral in the middle, which means that the quadrilateral is not of the dominance areas of any of the points, which is a contradiction.

For visual demonstration please visit <https://www.geogebra.org/geometry/dwmdtf5h>