Determining All Possible Separating Planes by Collision Tests

William Weimin Yoo

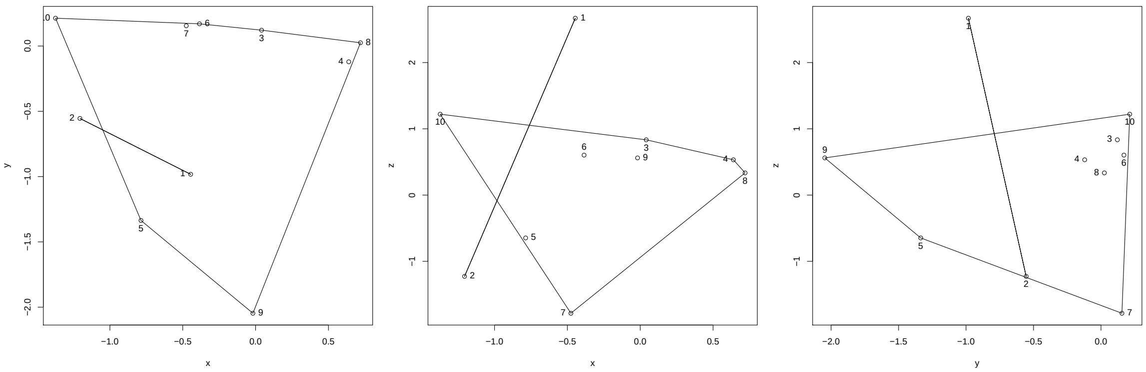
Abstract

Suppose we have points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ in $\mathbb{R}^3$. How many ways we can separate them using 2-dimensional planes? This is a classical problem in machine learning and the most commonly used approach to solving this problem is to recast it using integer programming. In this article, we propose another approach based on collision tests, where we conduct tests to see whether line segments between pair of points "collide" with the convex hull spanned by the other points.

Given a set of points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ in $\mathbb{R}^3$, we need to find the number of possible ways we can classify these points using two dimensional planes, where the classification is binary, i.e. 0 or 1. Hence, the total number of classification is 2^n in theory. However, based on the locations and configurations of these points, not all classification is possible. For $n \leq 2$, we can always shatter this set of points using planes, and hence we will have 2^n classifications in this case. For $n = 3$, we have two cases to consider. If these three points are not collinear, then we can always shatter these points with planes to get $2^3 = 8$ classifications. However, if there are collinear, we cannot use planes to separate the middle point. Since the middle point has two potential classifications, we will have only $2^3 - 2 = 6$ classifications. This is due to the fact that the Vapnik-Chervonenkis (VC) dimension of points in $\mathbb{R}$ is two with respect to lines. For the following, we say that a group of points is not linearly separable if they cannot be separated by planes

Now for $n \geq 4$, we need a fast and general way to detect all possible classifications. Here our basic strategy is to find combinations of points that are not linearly separable and subtract these cases from 2^n. For simplicity, let us illustrate the case for $n = 4$. We start by considering single points. We note that if a point is in the convex hull spanned by the rest three points, then this point cannot be separated by planes. To do this, we compute the convex hull of the four points and project it onto the x-y, x-z, and y-z planes respectively. We then check for interior points of each of the three projected convex hull. If a point is in all of the three convex hull interior, then this point is not linearly separable. Let us label this point as 1. Then this implies that the point combination 2-3-4 is also not separable. Therefore, we obtain this duality where if a certain group of points is not linearly separable, then its complement is also not separable. This further implies that for a set of points of length n, we only need to check point combination up to $\lfloor n/2 \rfloor$. Hence, in our present case, we need to check all single and double combinations of points only.

For set of two points, we will have $\binom{4}{2} = 6$ possible combinations. We start with the $1-2$ combination, where we have labeled the points 1 to 4 in a clockwise fashion. Then this $1-2$ combination is linearly separable iff the convex hull spanned by $1-2$ do not collide/intersect with the convex hull spanned by $3-4$. Both of these hulls are just lines in this case. To check for collision, we project these two hulls onto the x-y, x-z, and y-z planes respectively. Then at each of these planes, we again project these hulls onto the corresponding axis. For



(a) On x-y plane (b) On x-z plane (c) On y-z plane

Figure 1: Projected convex hulls for collision detection

example, after projecting the hulls $1-2$ and $3-4$ onto the x-y plane, we project them onto the x-axis and y-axis respectively. This then reduces to interval comparison in the axis. If the projected interval from the $1-2$ hull intersects with those from the $3-4$ hull in the x-axis, we say that a collision has occur in the x-axis of the x-y plane. Continuing in this fashion, we will have a total of six collision tests in each of the axis for each of the three planes. If collision is detected at each of the test, then we conclude that the convex hull $1-2$ and $3-4$ are not linearly separable. We repeat the same steps for the rest of the possible pairs.

Then the number of non linearly separable points m is twice the number of single non separable points plus the total number of pairwise non separable points. Then the number of possible classification is $2^4 - m$ in this case. Now, the same steps apply for $n \geq 5$. In the general case, two groups of points is linearly separable iff the convex hull spanned by the respective groups do not collide with each other. Collision is detected by projecting the three dimensional convex hull into one dimensional intervals, and collision occurs if the intervals intersect in all dimensions.

As an example, consider the case of ten points where each point is randomly generated by three i.i.d standard normal. Let say we want to determine whether points 1 and 2 can be separated by a plane from the rest. We see whether the projected convex hull spanned by $1-2$ collides with the projected convex hull spanned by the rest, in the x-y, x-z and y-z planes respectively. From Figure 1, we see that the projected convex hull (line) spanned by points $1-2$ collides with the projected convex hull spanned by the rest in all planes. Hence, this says that points $1-2$ is not linearly separable by planes from the rest.

We implemented our proposed collision method in the software package `R`. The program takes as input a $n \times 3$ matrix of numeric values, with each row representing each of the n points in $\mathbb{R}^3$; and outputs an integer representing number of possible classifications. The convex hull was constructed using the `chull` function from the base distribution.