Report on Convex Hull

Sabal Subedi
Master's in Computer Science
Idaho State University
Pocatello, ID, 83209 USA
sabalsubedi@isu.edu

February 10, 2024

Abstract

In this report, I will summarize the use of divide and conquer algorithm to obtain the convex hull of a set of points. I will give a pseudocode and analyze the asymptotic notation.

1 Introduction

The convex hull of a set of points is defined as the smallest convex polygon, that encloses all of the points in the set. Convex means that the polygon has no corner that is bent inwards.
The divide and conquer algorithm is a strategy of solving a large problem by breaking the problem into smaller sub-problems, solving the sub-problems and finally combining them to get the desired output. To implement this algorithm, we use recursion.
Here, we use divide and conquer algorithm to compute the convex hull of a set of points. First, I divide the set of points into equals halves recursively and then merge those halves using the logics that gives the final convex hull of the given set of points.

2 Pseudocode and Asymptotic Analysis

2.1 Pseudocode to compute convex hull

Algorithm 1 Convex Hull

```
 1: function CONVEXHULL($points$)
 2:     if $len(points) <= 3$ then
 3:         return $points$
 4:     $mid = len(points)//2$
 5:     $left\_hull$ ←CONVEXHULL($points[: mid]$)
 6:     $right\_hull$ ←CONVEXHULL($points[mid :]$)
 7:     $right\_most\_left\_hull$ ← maximum x-coordinate in left hull
 8:     $left\_most\_right\_hull$ ← minimum x-coordinate in right hull
 9:     $left\_decreasing = right\_decreasing = False$
10:     $left\_point, right\_point = right\_most\_left\_hull, left\_most\_right\_hull$
11:     $slope$ ← SLOPE($right\_most\_left\_hull, left\_most\_right\_hull$)
12:     while $left\_decreasing = False$ or $right\_decreasing = False$ do
13:         $left\_decreasing = right\_decreasing = True$
14:         while $left\_decreasing = True$ do
15:             $left\_point$ ← COUNTER_CLOCKWISE($left\_hull, left\_point$)
16:             $new\_slope$ ← SLOPE($left\_point, right\_point$)
17:             if $next\_slope < slope$ then
18:                 $slope = new\_slope$
19:             else
20:                 $left\_point$ ← CLOCKWISE($left\_hull, left\_point$)
21:                 $left\_decreasing = False$
22:             if SLOPE($left\_point$, CLOCKWISE($right\_hull, right\_point$)) $\leq$
    $slope$ then Break
```

```
        **while** $right\_decreasing = True$ **do**
            $right\_point \leftarrow$ CLOCKWISE($right\_hull, right\_point$)
            $new\_slope \leftarrow$ SLOPE($left\_point, right\_point$)
            **if** $next\_slope > slope$ **then**
                $slope = new\_slope$
            **else**
                $right\_point \leftarrow$ COUNTER_CLOCKWISE($right\_hull, right\_point$)
                $right\_decreasing = False$
            **if** SLOPE(COUNTER_CLOCKWISE($lelft\_hull, left\_point$), $right\_point$ ) $\geq$ $slope$ **then** Break
    $top\_left\_tan\_point, top\_left\_tan\_point = left\_point, right\_point$
    $left\_decreasing = right\_decreasing = False$
    $left\_point, right\_point = right\_most\_left\_hull, left\_most\_right\_hull$
    $slope \leftarrow$ SLOPE($right\_most\_left\_hull, left\_most\_right\_hull$)
    **while** $left\_decreasing = False$ or $right\_decreasing = False$ **do**
        $left\_decreasing = right\_decreasing = True$
        **while** $left\_decreasing = True$ **do**
            $left\_point \leftarrow$ CLOCKWISE($left\_hull, left\_point$)
            $new\_slope \leftarrow$ SLOPE($left\_point, right\_point$)
            **if** $next\_slope > slope$ **then**
                $slope = new\_slope$
            **else**
                $left\_point \leftarrow$ COUNTER_CLOCKWISE($left\_hull, left\_point$)
                $left\_decreasing = False$
            **if** SLOPE($left\_point$, COUNTER_CLOCKWISE($right\_hull, right\_point$)) $\geq$ $slope$ **then** Break
        **while** $right\_decreasing = True$ **do**
            $right\_point \leftarrow$ COUNTER_CLOCKWISE($right\_hull, right\_point$)
            $new\_slope \leftarrow$ SLOPE($left\_point, right\_point$)
            **if** $next\_slope < slope$ **then**
                $slope = new\_slope$
            **else**
                $right\_point \leftarrow$ CLOCKWISE($right\_hull, right\_point$)
                $right\_decreasing = False$
            **if** SLOPE(CLOCKWISE($lelft\_hull, left\_point$), $right\_point$ ) $\leq$ $slope$ **then** Break
    $lower\_left\_tan\_point = left\_point$
    $lower\_right\_tan\_point = right\_point$
    $final\_hull = []$
    $right\_hull\_point = top\_right\_tan\_point$
    while $right\_hull\_point \neq lower\_right\_tan\_point$ do
        $final\_hull.APPEND(right\_hull\_point)$
        $right\_hull\_point \leftarrow$ CLOCKWISE$(right\_hull, right\_point)$
    $final\_hull.APPEND(lower\_right\_tan\_point)$
    $left\_hull\_point = lower\_left\_tan\_point$
    while $left\_hull\_point \neq top\_left\_tan\_point$ do
        $final\_hull.APPEND(left\_hull\_point)$
        $left\_hull\_point \leftarrow$ CLOCKWISE$(left\_hull, left\_point)$
    $final\_hull.APPEND(lower\_left\_tan\_point)$
    return $final\_hull$
```

Time complexity and space complexity
Given,
points: a list of tuples(x,y)

Space Complexity:

1. Since the depth of recusion for this algorithm is $O(\log n)$ space.
2. To store variables like left_hull, right_hull, final_hull is atmost the size of input. So, it takes $O(n)$ space.

$$S(n) = O(\log n) + O(n) \approx O(n) \tag{1}$$

Time Complexity:

1. The recursive division of a set of points into halves talkes $T(n) = 2T(n/2) + O(n)$. Using Master Theorem, we get $O(n \log n)$
2. To find the left hull and right hull, i.e. dividing the points at each recursion, takes $O(n)$
3. To find the upper tangent, the while loop runs $O(n)$ times
4. To find the lower tangent, the while loop runs $O(n)$ times

5. Finally, to merge the hull points it takes $O(n)$ time

$$\begin{aligned} T(n) &= O(n \log n) + O(n) + O(n) + O(n) + O(n) \\ &\approx O(n * \log n) \end{aligned} \tag{2}$$

Analysis of algorithm:
The pseudocode takes set of points and outputs a single set. It uses divide and conquer algorithm to divide the given set of points into two halves recursively. Finally, it merges the hull points to get the final convex hull.

2.2 Pseudocode to get next point based on clockwise order

Algorithm 2 Get points on Clockwise order

```
1: function CLOCKWISE($points, point$)
2:     $index \leftarrow$ index of the point in points
3:     if $index = len(points) - 1$ then
4:         return $points[0]$
5:     else
6:         return $points[index + 1]$
```

Time complexity and space complexity
Given,
points: a set of points
point: a valid tuple(x,y)
Space Complexity:

$$S(n) = O(1) + O(1) + O(1) = O(3) \approx O(1) \tag{3}$$

Time Complexity:

$$\begin{aligned} S(n) &= O(1) + O(1) + O(1) + O(1) + O(1) + O(1) \\ &= O(6) \approx O(1) \end{aligned} \tag{4}$$

Analysis of algorithm:
The pseudocode takes two inputs: a set of points and a valid tuple(x,y) and outputs a tuple(x,y). It returns next point in clockwise order of given point in the set of points.

2.3 Pseudocode to get next point based on counterclockwise order

Algorithm 3 Get points on Counter-Clockwise order

```
1: function COUNTER_CLOCKWISE($points, point$)
2:     $index \leftarrow$ index of the point in points
3:     if $index = 0$ then
4:         return $points[len(points) - 1]$
5:     else
6:         return $points[index - 1]$
```

Time complexity and space complexity
Given,
points: a set of points
point: a valid tuple(x,y)
Space Complexity:

$$S(n) = O(1) + O(1) + O(1) = O(3) \approx O(1) \tag{5}$$

Time Complexity:

$$\begin{aligned} S(n) &= O(1) + O(1) + O(1) + O(1) + O(1) + O(1) \\ &= O(6) \approx O(1) \end{aligned} \tag{6}$$

Analysis of algorithm:
The pseudocode takes two inputs: a set of points and a valid tuple(x,y) and outputs a tuple(x,y). It returns next point in counter-clockwise order of given point in the set of points.

2.4 Pseudocode to compute slope

Algorithm 4 Slope

```
1: function SLOPE($point1, point2$)
2:     $x_1, y_1, x_2, y_2 = point1[0], point1[1], point2[0], point2[1]$
3:     return $((point2[1] - point1[1])/(point2[0] - point1[0]))$
```

Time complexity and space complexity
Given,
point1: a valid tuple(x,y)
point2: a valid tuple(x,y)
Space Complexity:

$$S(n) = O(1) + O(1) + O(1) + O(1) + O(1) = O(5) \approx O(1) \tag{7}$$

Time Complexity:

$$T(n) = O(1) + O(1) \approx O(1) \tag{8}$$

Analysis of algorithm:
The pseudocode takes two points as input and outputs a single value. It computes the slope between two points and return the result.

3 Observations and Results

3.1 Observation Table

Sample Size	Distribution	Time
10	Gaussian	0.000 sec
100	Gaussian	0.001 sec
1000	Gaussian	0.004 sec
10000	Gaussian	0.056 sec
100000	Gaussian	0.667 sec
500000	Gaussian	2.983 sec
1000000	Gaussian	5.502 sec

Here, I have chosen Gaussian distribution to collect the elapsed time (i.e. time taken to get the convex hull) of seven different sample size. The elapsed time increases as the size of the sample increases.
To give more insight on this, I have collected 5 different samples for each sample size and get the data to compute the mean time needed for each sample size. Using the obtained data, I have plotted a graph shown in figure 1.

.

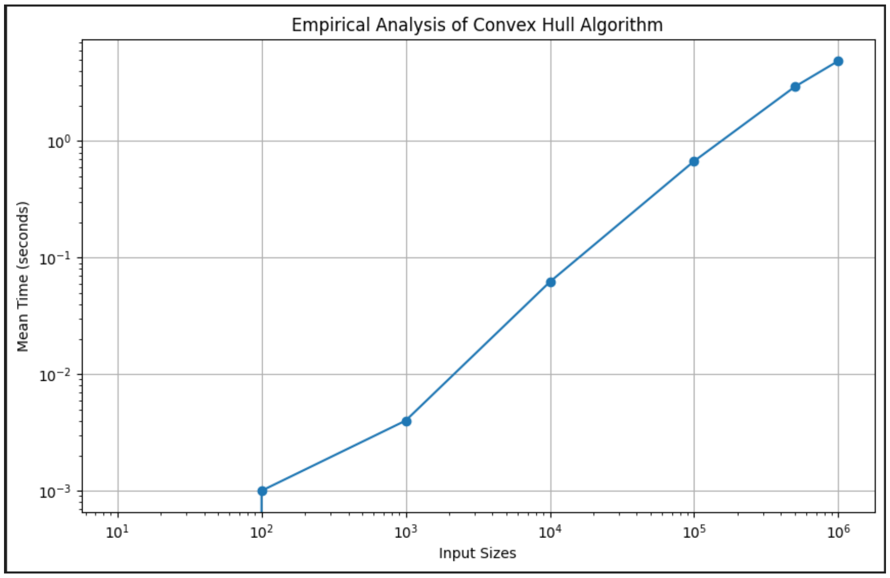


Figure 1: sample size vs mean time required

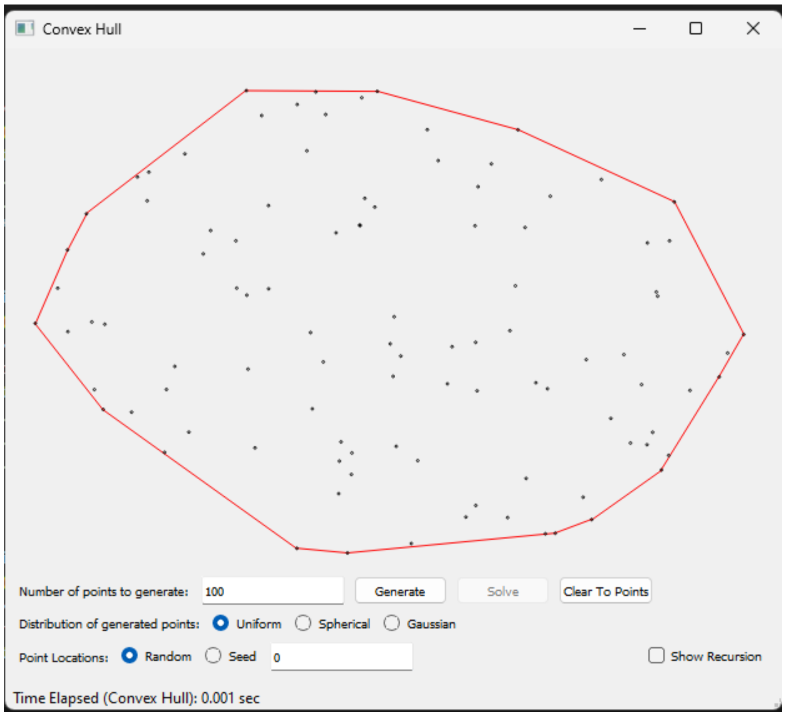
In the graph (in figure 1), we can see a increasing curve i.e. the mean time required to compute the convex hull with smaller sample size takes less time and so on. The change in size and the mean time follows logarithmic pattern. This mean that the changes either increasing or decreasing in sample size changes the mean time by increasing or decreasing rate.
According to my analysis on time complexity of CONVEXHULL algorithm, it takes $O(n * \log n)$ time. But, my graph (empirical) analysis differs from the theoretical analysis. The graph looks identical with $1/3(n * \log n)$. Therefore, the constant of proportionality(k) is $1/3$ so that $CH(Q) = K * g(n)$

Reasons for the difference between theoretical and empirical analysis:

1. The hardware capacity of the machine can be a reason.
2. The sparseness of the data can be a valid reason.
3. Theoretical analysis focus on algorithm efficiency and estimates its performance based on input size and its growth. However, this can differ on empirical analysis. Implementation details, hardware limitations, etc can influence the empirical analysis.

3.2 Results

.



Number of points to generate: 100

Distribution of generated points: [✓ Uniform] Spherical Gaussian

Point Locations: [✓ Random] Seed 0

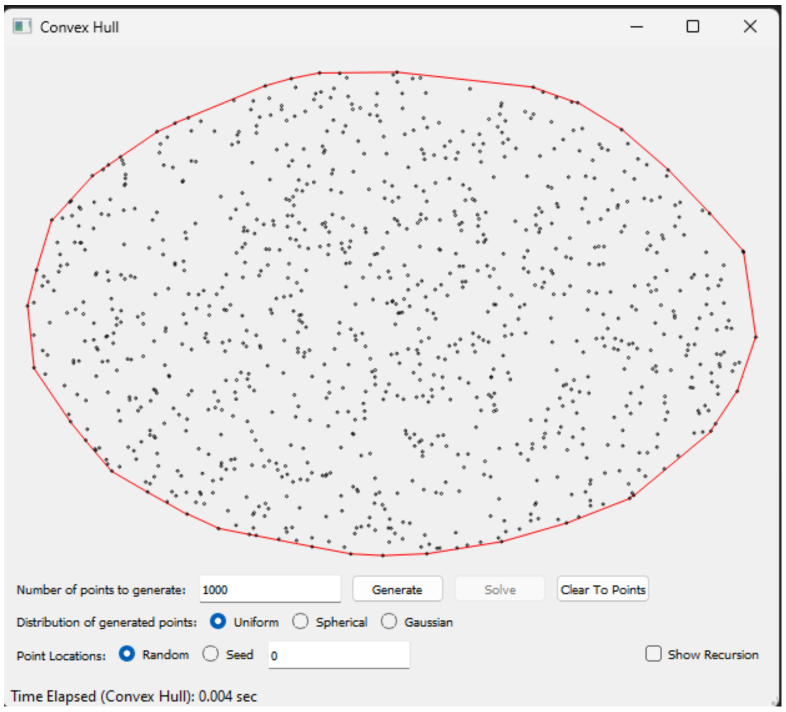
Time Elapsed (Convex Hull): 0.001 sec

Figure 2: Convex hull for 100 points

I used 100 sample size (in figure 2) to compute hull. As expected the program was able to compute and display convex hull. The program took 0.001 second to compute the convex hull under Uniform distribution and random seed.

Again, I used 1000 random points to test the program under same setup and the program was able to generate the convex hull as expected. It took 0.010 seconds. The change in elpased time (required time) is significant as the sample size increases.

.



Number of points to generate: 1000

Distribution of generated points: [✓ Uniform] Spherical Gaussian

Point Locations: [✓ Random] Seed 0

Show Recursion

Time Elapsed (Convex Hull): 0.004 sec

Figure 3: Convex hull for 1000 points

4 Conclusion

In summary, I was able to compute a convex hull using divide and conquer algorithm along with orientation and distance. I was able to analyze the pseudocode and get the time and space complexity.