

Lecture 6: January 23, 2018

Lecturer: Bin Ma

Notes By: Harsh Mistry

6.1 Divide and Conquer Continued

Example 6.1 Closest Pair

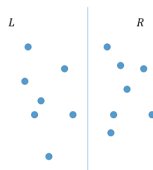
Input : A set of points on plain: $P = \{p_1, \dots, p_n\}$ Each $p_i = (x_i, y_i)$. Euclidean distance.

Output : A pair of points p_1, p_j that are closest among all pairs

We use divide and conquer. The algorithm we are presenting needs the points to be sorted. Supposed P are sorted with x axis. Additionally, we have P' containing references to all the points and sorted according to y -axis. This requires $O(\log m)$ time to sort the points. So we have P, P' to continue.

For simplicity we assume all points have different x -axis for now.

First, we split P into two parts L and R . With the assistance of I_x , this can be done efficiently.



before we claim we can recursively solve the left and right parts, there is one more thing to do. Since we assume the bigger problem has P' that are sorted according to y -axis, we need also ensure that the two smaller problems also have the y -axis sorted reference array. What we can do is to linearly scan through P' , and put the points into L' and R' , depending on whether the point is at left or right of the split line. This takes $O(n)$ time.

Now we can solve the left and right parts recursively. Let the least distance be γ_1 and γ_2 . Let $\gamma = \min\{\gamma_1, \gamma_2\}$. The only remaining problem is the closest pair may be across the two parts nearby the split line.

If $l \in L$ and $r \in R$ are such that $d(l, r) \leq \gamma$, then they must be within distance γ of the split line. So, we only need to limit our search within a narrow band around the split line as shown in the figure. We denote the list of points in this band by S . Since P is sorted according to x -axis, S can be found in linear time. Additionally, we can linearly scan through P' to construct an array S' of references to points of S , and make S' sorted according to y -axis

Our goal is to find a pair (p, q) from S within distance γ . For each point $p \in S$ with y -axis p, y , we only need to compare it with the other points with y -axis between p, y and $p, y + \gamma$. The above analysis indicates that there are at most 11 such points in S . They are all clustered together in the y -axis sorted S' . The search takes $11|S| = O(n)$ time. Let the minimum distance we find is γ_m . Finally, we output $\min\{\gamma, \gamma_m\}$. If desired, the pair with the shortest distance can also be recorded during the above computation for output

Time Complexity :

$$T(n) = 2 \cdot T\left(\frac{n}{2}\right) + O(n) \implies T(n) = O(n \log n)$$