Lecture 3: Convex Hull

Jingze Cui

September 25,2017

Big data Algorithm:

$$\text{Relationship among Data-Computer}\begin{cases} size\quad of\quad memory\quad handdisk \\ speed\quad sub-linear\quad time\quad difficulty \\ compression? --space \\ Kemelization:\quad difficult\quad task\quad focused\quad on\quad a\quad kernel \\ Output\quad sensitive\begin{cases} if\quad output\quad is\quad small \\ algorithm\quad work\quad falter \end{cases} \end{cases}$$

Output sensitive Algorithm: Use structure within data, so that we do things faster.

1 convex hull

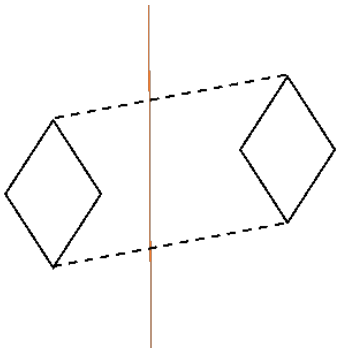

Figure 1

Input: n points in the plane.$p_1, p_2, ..., p_n \qquad p_i = [x_i, y_i]$

Divide & Merge

(Merge sort: Divide, sort both points, merge two sorted list.) Sort $(x_i : 1 \leq i \leq n) \quad x_1 < x_2 < ... < x_n$

Divide $P = P_L \bigcup P_R \begin{cases} P_L = \{P_i : 1 \leq i \leq \frac{n}{2}\} \\ P_R = P - P_L \end{cases}$

Conquer $C_L = Convex(P_L), \quad Convex(P_R)$

Merge: carefully done, in linear time

$\begin{cases} \quad min \quad b \\ s.t. \quad \forall i, \quad ax_i + b \geq y_i \end{cases}$ 2 variable find tangent of n points in the plane

Solve the tangent problem to find a,b is in linear time $T = o(n)$

$T(n) = 2 \times T(\frac{n}{2}) + o(n) \quad \Rightarrow T(n) = o(n \log n)$

where $T(\frac{n}{2})$means divide the points into two parts, $o(n)$ means the time spend to find tangent.

Randomly pick points from unit square. How many points are in the convex hull? Expected number of points on convex hull → o(1) Is there a fast algorithm to find contex hull if the number of points on the convex hull is small?

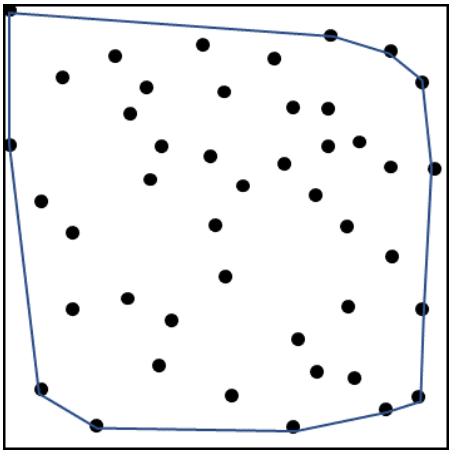

Figure 2

Alternatively split points into two parts, find their common targets, throw away a lot of points.

For $x_1, x_2, ..., x_n$, find the medium of $\{x_i, 1 \leq x_i \leq n\}$

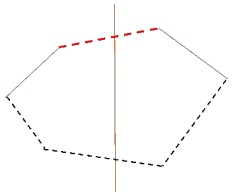

Figure 3

h = number of edges on up hull

prove $T(n) = c * (n \log h) = o(n \log h)$

Find mediam of $\{x_i : 1 \leq i \leq n\}$

min $ax_{median} + b \geq y_i, for \quad i = 1, 2, ..., n$

$T(n, h) = T(\frac{n}{2}, h_1) + T(\frac{n}{2}, h_2) + n$

$h_1 + h_2 + 1 = h$, h_1 means the left part of up hull, h_2 means the right part of up hull, 1 means the tangent of two part.

By induction: $T(\frac{n}{2}, h_1) = \frac{cn}{2} \log h_1$, $T(\frac{n}{2}, h_2) = \frac{cn}{2} \log h_2$

$$\begin{aligned} T(n, h) &= \frac{cn}{2} \log h_1 h_2 + n \\ &\leq \frac{cn}{2} \log (\frac{h}{2})^2 + n \\ &= cn(\log h - 1) + n \end{aligned}$$

when $c \geq 1$, $T(n, h) = o(n \log h)$