

一般选择问题 的算法设计

一般性选择问题

问题：选第 k 小.

输入：数组 S , S 的长度 n , 正整数 k ,
 $1 \leq k \leq n$.

输出：第 k 小的数

实例 1

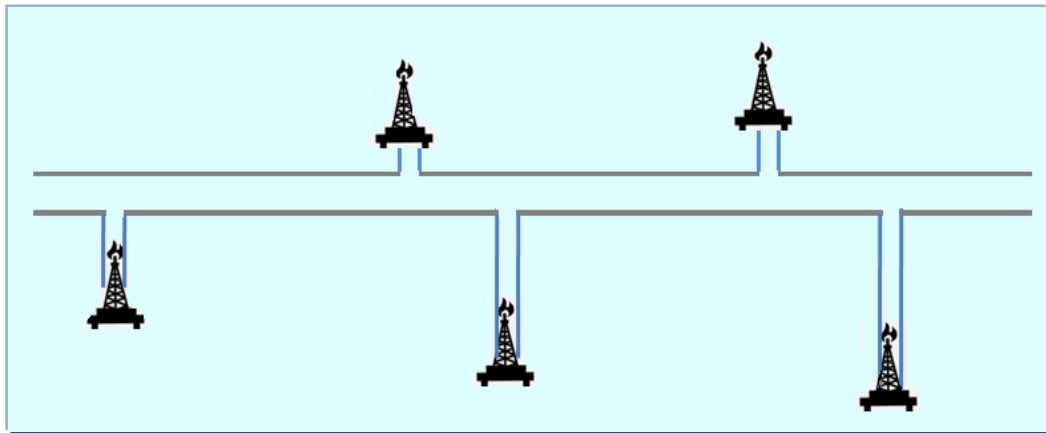
$S = \{ 3, 4, 8, 2, 5, 9, 18 \}$, $k = 4$, 解: 5

实例 2

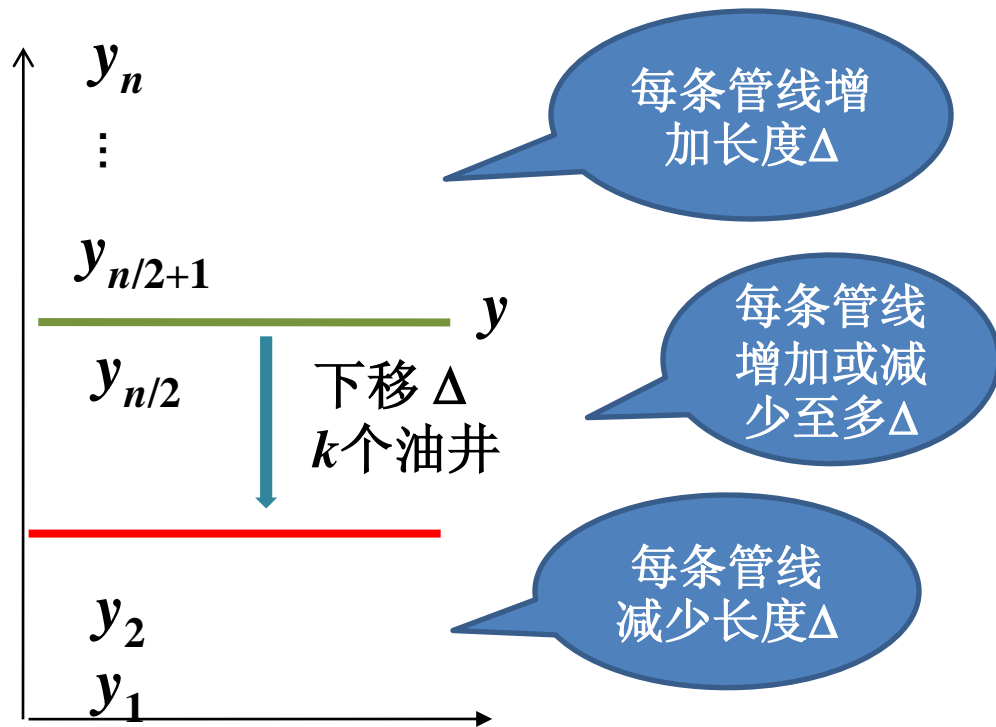
统计数据的集合 S , $|S| = n$,
选中位数, $k = \lceil n/2 \rceil$

一个应用：管道位置

问题：某区域有 n 口油井，需要修建输油管道. 根据设计要求，水平方向有一条主管道，每口油井修一条垂直方向的支管道通向主管道. 如何选择主管道的位置，以使得支管道长度的总和最小？



最优解: Y 坐标的中位数



下移后支管线总长度增加

简单的算法

算法一：

调用 k 次选最小算法

时间复杂度为 $O(kn)$

算法二：

先排序，然后输出第 k 小的数

时间复杂度为 $O(n \log n)$

分治算法

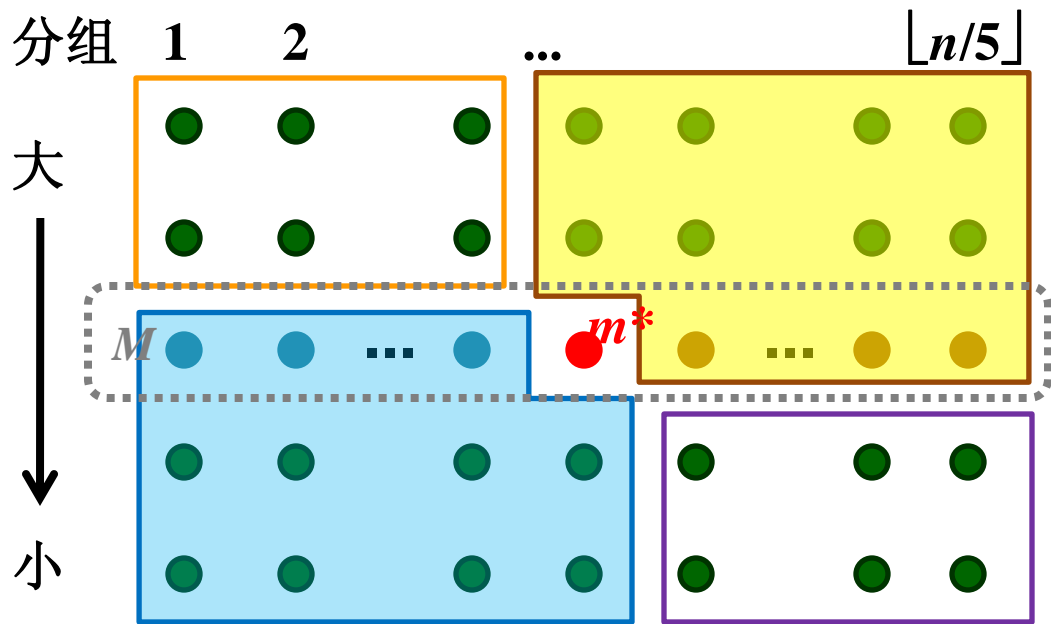
假设元素彼此不等，设计思想：

1. 用**某个元素 m^*** 作为标准将 S 划分成 S_1 与 S_2 ，其中 S_1 的元素小于 m^* ， S_2 的元素大于等于 m^* 。
2. 如果 $k \leq |S_1|$ ，则在 S_1 中找第 k 小。
如果 $k = |S_1| + 1$ ，则 m^* 是第 k 小
如果 $k > |S_1| + 1$ ，则在 S_2 中找第 $k - |S_1| - 1$ 小



算法效率取决于子问题规模，
如何通过 m^* 控制子问题规模？

m^* 的选择与划分过程



A: 数需要与 m^* 比大小, **B**: 数大于 m^*

C: 数小于 m^* , **D**: 数需要与 m^* 比大小

实例: $n=15, k=6$

8	2	3	5	7	6	11	14	1	9	13	10	4	12	15
---	---	---	---	---	---	----	----	---	---	----	----	---	----	----

	8	14	15	
	7	11	13	
M	5	9	12	$m^* = 9$
	3	6	10	
	2	1	4	

	8	14	15	
A	7	11	13	B
	5	9	12	
C	3	6	10	D
	2	1	4	

8, 7, 10, 4 需要与9比较

归约为子问题

S_1	8	14	15	S_2
	7	11	13	
	5	9	12	
	3	6	10	
	2	1	4	

子问题

8 7 5 3 2 6 1 4

子问题规模 = 8, $k=6$

伪码

算法 Select (S, k)

输入：数组 S ，正整数 k ，

输出： S 中的第 k 小元素

1. 将 S 分 5 个一组，共 $n_M = \lceil n/5 \rceil$ 组
2. 每组排序，中位数放到集合 M
3. $m^* \leftarrow \text{Select}(M, \lceil |M|/2 \rceil)$ // S 分 A, B, C, D
4. A, D 元素小于 m^* 放 S_1 , 大于 m^* 放 S_2
5. $S_1 \leftarrow S_1 \cup C; S_2 \leftarrow S_2 \cup B$ 划分
6. if $k = |S_1| + 1$ then 输出 m^*
7. else if $k \leq |S_1|$ \Leftarrow 递归计算子问题
8. then Select (S_1, k)
9. else Select ($S_2, k - |S_1| - 1$)

小结

选第 k 小的算法:

- 分治策略
- 确定 m^*
- 用 m^* 划分数组归约为子问题
- 递归实现