

# k 分位数

2020 年 11 月 8 日

## 1 期中考试样卷 2-3

### 题目描述：

对一个包含  $n$  个元素的集合， $k$  分位数是指能把有序集合分成  $k$  个等大小集合的第  $k-1$  个顺序统计量。给出一个能找出某一集合的  $k$  分位数的  $O(n \lg k)$  时间的算法。

The  $k$ th quantiles of an  $n$ -element set are the  $k-1$  order statistics that divide the sorted set into  $k$  equal-sized sets (to within 1). Give an  $O(n \lg k)$ -time algorithm to find the  $k$ th quantiles of a set.

### 思路分析：

算法思想：分治策略，先找  $k$  分位数中的第  $\lfloor k/2 \rfloor$  位数，然后以此为主元将数组分为两部分，不断递归处理。

子算法：SELECT 算法，因为其时间复杂度为  $O(n)$ 。

时间复杂度分析：不妨假设  $n$  和  $k$  都为 2 的指数。用  $T(n)$  表示问题规模为  $n$  时算法的运行时间，则可以写出递归式为  $T(n) = 2T(n/2) + cn$ ， $c$  为某个常数。且  $T(n/k) = O(1)$ 。则：

$$\begin{aligned} T(n) &\leq cn + 2T(n/2) \\ &\leq cn + 2c(n/2) + 4T(n/4) \\ &= 2cn + 4T(n/4) \\ &\leq 3cn + 8T(n/8) \\ &\dots \\ &\leq \log(k) \, cn + kT(n/k) \\ &\leq \log(k) \, cn + O(k) \\ &= O(n \log k) \end{aligned}$$