

数据结构与算法I 思考题4

2019201409 于倬浩

问题

两个排好序的数组，长度分别为 n 和 m ，求这 $n+m$ 个数的中位数。

做法

不妨设有两个数组 $a[1 \dots n]$ ， $b[1 \dots m]$ 。如果 n 大于 m ，显然可以直接交换两个数组，保证 n 小于等于 m 。

接下来，先考虑 $n+m$ 为奇数的情况：此时中位数唯一存在，即为第 $\lceil \frac{n+m}{2} \rceil$ 个元素，设 $T = \lceil \frac{n+m}{2} \rceil$ 。

那么，我们可以在较小的数组 a 上二分，假设当前二分到的元素下标为 p ，那么如果想判断 p 是否为中位数，还需在 b 中找到下标 $q = T - p$ 的元素。接下来：

- 如果 $q < m$ 且 $b[q] \leq a[p] \leq b[q + 1]$ ，或者 $q = m$ 且有 $b[q] \leq a[p]$ ，则说明 $a[p]$ 恰好为合并后的数组的排名为 T 的元素，即 $a[p]$ 即为我们想要的中位数。
- 如果 $p < n$ 且 $a[p] \leq b[q] \leq a[p + 1]$ ，或者 $p = n$ 且有 $a[p] \leq b[q]$ ，则说明 $b[q]$ 恰好为合并后数组排名为 T 的元素，即 $b[q]$ 即为我们想要的中位数。
- 如果 $a[p] < b[q - 1] \leq b[q]$ ，那么至少 $a[p]$ 、 $b[q]$ 都不可能成为中位数，且最终中位数对应的 a 中下标一定大于 p ，因此将二分的左端点设为 $p+1$ 。

- 如果 $a[p] > b[q + 1]$ ，那么至少 $a[p]$ 、 $b[q]$ 都不可能成为中位数，且最终中位数对应 a 中下标一定小于 p ，因此二分端点设为 $p-1$ 。

对于数组长度为偶数的情况，中位数为排名为 $\frac{n+m}{2}$ 和排名为 $\frac{n+m}{2} + 1$ 两个元素的平均数，因此套用上述代码，找到排名为 T 和 $T + 1$ 的两个元素，直接求平均即可。