

区间众数解题报告

杭州外国语学校 陈立杰

Contents

1	题目大意	2
2	文中会用到的一些记号说明	2
3	只有询问的情况	3
3.1	定理1	3
3.2	定理1的证明	3
3.3	算法	3
3.4	做法1的优化	3
4	考虑修改操作	5
4.1	算法	5

www.docin.com

1 题目大意

给定长度 n 的数列 a ,每个数在1到 n 之间,要求支持一些操作

- 1 修改第 i 个数的值
- 2 询问区间 $[l, r]$ 中,出现次数最多的数的出现次数.

操作一共有 q 个

数据满足 $n, q \leq 50000$

2 文中会用到的一些记号说明

集合都是多重集合,既一个数可以出现多次.

众数就是出现次数最多的数.

一个集合 a 的众数 $\text{mode}(a)$.



3 只有询问的情况

修改操作的出现对解题造成了很大影响, 我们不妨先考虑没有这些操作的情况.

3.1 定理1

那么考虑集合 a, b , 那么 $\text{mode}(a \cup b) \in \text{mode}(a) \cup b$

3.2 定理1的证明

这个几乎是显然的, 因为如果 t 既不是 $\text{mode}(a)$ 也不属于 b , 那么它在 $a \cup b$ 中的出现次数就是在 a 中的出现次数, 而这不会比 $\text{mode}(a)$ 出现的更多.

3.3 算法

我们不妨将 n 个数分成 \sqrt{n} 块, 每块大小 $O(\sqrt{n})$. 并且预处理出数组 $a[i][j]$ 表示第 i 块到第 j 块这些数的众数.

那么考虑询问 $[l, r]$. 如果区间 $[l, r]$ 在某一个块内部, 那么我们暴力得出答案, 复杂度为 \sqrt{N} . 不然令 l 在第 a 块, r 在第 b 块. 那么 $[l, r]$ 就由

- 1 l 到第 a 块最后一个
- 2 第 $a + 1$ 块到第 $b - 1$ 块
- 3 第 b 块第一个到 r

这3部分组成. 注意到第2部分的答案我们已经处理出来了, 那么根据定理1, $[l, r]$ 的众数要么是第2部分的众数, 要么是第1部分或者第3部分中的数. 由于第1, 3部分的大小最大是 $2\sqrt{n}$, 我们只需要枚举这些数判断它们在 $[l, r]$ 中的出现次数即可.

我们只需要对每种不同的数, 对他们的位置二分查找, 就能回答: $[l, r]$ 中有几个 x 这样的问题.

那么一次询问复杂度就是 $O(\sqrt{n} \log n)$.

同时预处理只要用类似方法就能做到 $O(n^{1.5} \log n)$

那么总时间复杂度就是 $(n + q)\sqrt{n} \log n$

3.4 做法1的优化

由于算法的本质是分块, 所以 \sqrt{n} 这个系数是没法去掉了. 不过我们来考虑如何去掉 $\log n$ 这个系数.

要去掉 $\log n$ 这个系数, 我们需要在 $O(1)$ 时间内回答 $[l, r]$ 之间有几个 x 这样的问题.

不妨令 $F(i, x)$ 表示 $[0, i]$ 间有几个 x .

那么 $[l, r]$ 间的 x 数量就是 $F(r, x) - F(l - 1, x)$.

我们只需要 $O(1)$ 回答 $F(i, x)$

不妨将 n 个数分成 $O(\sqrt{n})$ 块,每块大小 $\lfloor \sqrt{n} \rfloor$ 预处理出 $C[i][x]$ 表示第0块到第 i 块中, x 出现了多少次. 这个需要时间和空间 $O(n^{1.5})$

同时对于每个块 b ,预处理 $A[b][i][x]$ 出它的前 i 个中 x 出现了多少次.注意到这里的 x 最多只有 \sqrt{n} 个那么每个预处理每个块需要 $(\sqrt{n})^2 = n$ 的时间和空间.

我们需要得到值 x 在第 i 个块内部的id,可以对每个块开一个表来实现.

总预处理时间就是 $O(n^{1.5})$

询问也只需要分成前几块内部和最后一个块前几个2部分计算即可.

那么优化后的时间复杂度就是 $((n + q)\sqrt{n})$.



4 考虑修改操作

4.1 算法

由于有了修改操作,我们之前的做法就需要修改了.

考虑我们讲数列分成大小基本相等的 L 块,一次修改可能很改变 $O(L^2)$ 个块之间的结果,如果还取 $L = \sqrt{n}$,那么每次修改就会至少需要 $O(L^2) = O(n)$ 的时间,跟朴素无异了.

不妨令 $L = \sqrt[3]{n}$.那么我们维护 $O(L^2)$ 对块之间的结果.

我们对每对块之间,维护每个值在其中出现了多少次和最大出现次数. 同时用 $cnt[i]$ 表示出现了 i 次的值有几个.

如果一次将 a 到 b 块之间的一个数修改了,那么一个值出现次数 -1 ,一个值出现次数 $+1$.

2个操作都最多将最大出现次数改变1,判断一下完成修改即可.

同时考虑回答询问 $F(i, x)$.

考虑之间的预处理方法,如果需要修改的话,经过分析可以发现需要修改 $O(\sqrt{n})$ 个数.不影响整体复杂度.

那么我们就能够 $O(1)$ 回答 $[l, r]$ 之间有几个 x 这个询问.

那么修改复杂度就是 $O(L^2) = O(n^{2/3})$. 询问复杂度就是 $O(n/L) = O(n^{2/3})$. 预处理复杂度是 $O(n^{5/3})$.

总时间复杂度就是 $O((n + q)n^{2/3})$