2012 集训队互测 世博会(刘洪轩) 解题报告

天津市南开中学
刘洪轩

一、题意简述

有 N 个物品，编号从 1 到 N，每个物品有两个属性 Ai 和 Bi，我们定义两个物品的差别为它们对应属性差的绝对值中较大的一个。之后有 Q 个询问，对于每个询问，给出 Li 和 Ri，然后需要我们要设定一个属性为 A'和 B'的虚拟物品，使这个虚拟物品与从 Li 到 Ri 的每一个物品的差异度之和最小，然后对每个询问输出最小的差异度数值。

$1<=N<=100000, 1<=Q<=100000, |Ai|,|Bi|<=1000000000$

二、题目分析

首先对题目条件进行转化，若两个物品 A1,B1 和 A2,B2，它们的差异为

$$
\begin{aligned}
&MAX\{|A1-A2|, |B1-B2|\}\\
&= MAX\{A1-A2, B1-B2, B2-B1, A2-A1\}\\
&= MAX\{(A1+B1)/2 - (A2+B2)/2) + (A1-B1)/2 - (A2-B2)/2,\\
&\qquad (A1+B1)/2 - (A2+B2)/2) + (A2-B2)/2 - (A1-B1)/2,\\
&\qquad (A2+B2)/2 - (A1+B1)/2) + (A1-B1)/2 - (A2-B2)/2,\\
&\qquad (A2+B2)/2 - (A1+B1)/2) + (A2-B2)/2 - (A1-B1)/2\}\\
&= MAX\{(A1+B1)/2 - (A2+B2)/2, (A2+B2)/2 - (A1+B1)/2\}\\
&\quad + MAX\{(A1-B1)/2 - (A2-B2)/2, (A2-B2)/2 - (A1-B1)/2\}\\
&= |(A1+B1)/2 - (A2+B2)/2| + |(A1-B1)/2 - (A2-B2)/2|
\end{aligned}
$$

(转化过程中利用了一条重要的性质 $|x| = MAX\{x, -x\}$)

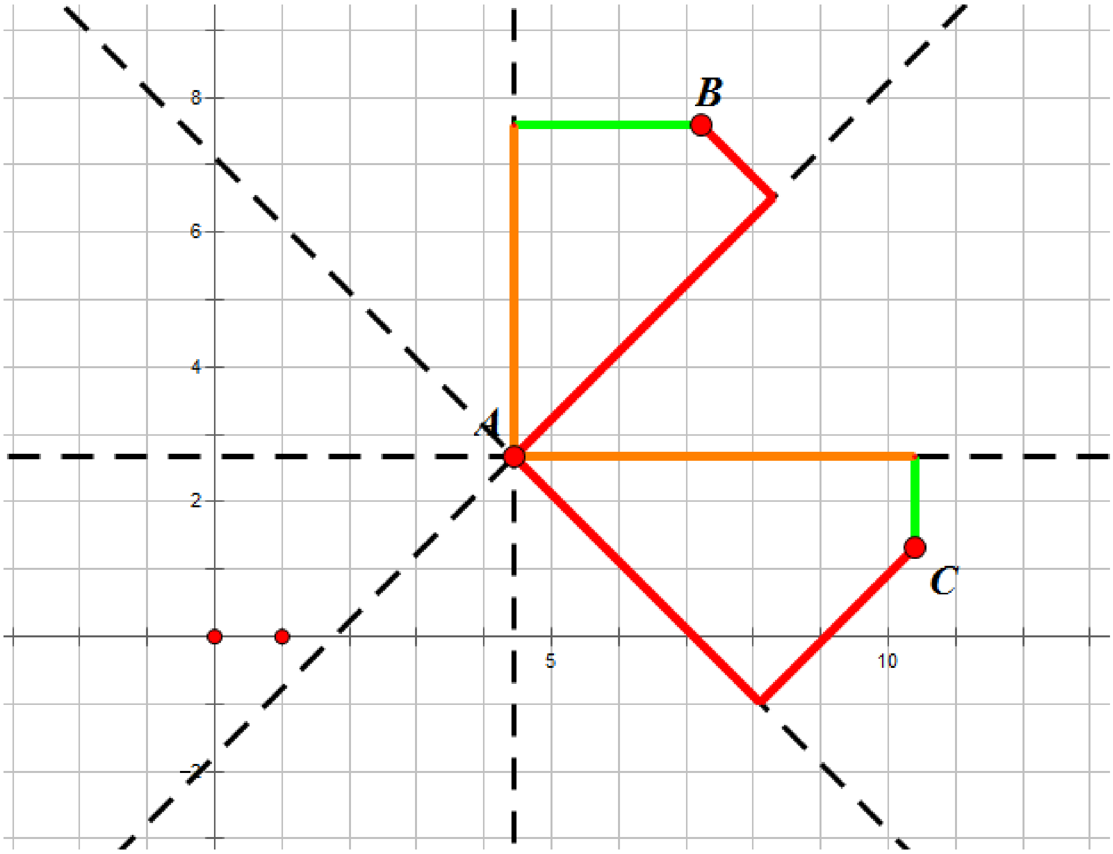
然后我们令 $Xi = (Ai+Bi)/2$, $Yi = (Ai-Bi)/2$，并让每一个物品对应平面上坐标为(Xi, Yi)的点。这样两个物品的差异就变成了$|X1-X2|+|Y1-Y2|$，即它们在平面对应的点的曼哈顿距离。

而我们要求的是和一些物品差异度之和最小的物品，我们把其转化为求距离一些点曼哈顿距离之和最小的点。当我们求出这个点且坐标为(X', Y')时，我们所需的物品的两个属性 A'和 B'就可以表示为

$A' = X' + Y', B' = X'- Y'$

转化为平面图上的曼哈顿距离后，题目就显得简单了。

而上文中的推导步骤看着又长又繁琐，但这个转化还是比较容易想到的。我们若以物品的两个属性 A 和 B 为坐标在平面上取点，我们会发现，两个点的两个坐标差较大的一个和这两个点在斜向方向上的曼哈顿距离成比例，如下图所示



A 到 B 和 A 到 C 的“差异”都转化为了斜向的曼哈顿距离。观察出这个之后，上文中的推导得出的结论也不难发现。

之后我们来解决最小曼哈顿距离和的问题。我们可以看到，曼哈顿距离在 x 轴方向和 y 轴方向是独立的，我们可以把其拆开考虑，如下面的推导所示

$\sum(|X_i-X'|+|Y_i-Y'|) = \sum|X_i-X'| + \sum|Y_i-Y'|$

我们把问题拆成了两个子问题之和，这个子问题是求到数轴上多个点距离之和最小的点。很明显的，这个点的坐标应该和把所有点排序后处于正中央的那个点坐标相同。若这多个点的数目是偶数，则坐标取中间的两个点之间的任意一个位置都可以。

这个结论证明也很简单，我们把所以点两两配对，最左边和最右边的配成一对，左边数第二个和右边数第二个配成一对。我们取的点到某一对点的距离和总是大于等于这一对点的距离且我们取的点在这一对点之间时取等。所以我们要把我们取的点放在每一对点之间，即所有点的中位数的位置。

这样的话，我们计算出所有的 Xi 和 Yi，然后对应每一个询问，求出对应范围 X 的中位数和 Y 的中位数作为 X’和 Y’。然后可以求出 A’和 B’然后计算每一个差异度再求和，也可以用 X’和 Y’计算出每一个曼哈顿距离再求和，这两种方法都能得到最后要输出的答案。

若我们用快速查找的方法求中位数，这个方法的复杂度就是 O(qn)，明显不满足题目的要求。而实际上，我们求中位数的时候可以利用一些高级的数据结构，如划分树可以在 O(nlogn)-O(logn)的时间内求出区间第 k 大数，并且也能在求出第 k 大数的同时求出我们所需要的曼哈顿距离和。使用划分树可以通过这道题的所有测试点，我们后面也主要讲解一下划分树的大体实现方法。

三、划分树的实现

建立划分树之前需要把序列中的所有数进行稳定的排序后生成一个新的序列。

划分树的结构是基于线段树的，根节点包含原来整个序列，每个节点包含一段序列，叶节点包含一个长度为 1 的序列。每个节点包含的范围对应的是排序后的序列中的范围，就是若一个节点对应的线段是 l 到 r，它就包含第 l 小到第 r 小的所有数字。但这些数字在每个节点中的顺序不是按大小关系排列的，而是按在原序列中的顺序排列的。

建立划分树时，根节点就包含原来的整个序列。之后建立每一个节点的左右子树时，先从排序后的序列中查询出这个节点的包含的数字的中位数，然后从左往后扫描每一个数字，小于等于中位数的放到左子树中，大于中位数的放到右子树中。并且每处理一个数字就计算出到这里为止对应左子树的哪一个位置。记录的话我们可以把节点中的每一个子序列对应的左子树和右子树中的子序列求出来。构建直到一个节点只包含 1 个数为止。可以看到划分树有 O(logn)层，且构建划分树的时间复杂度以及划分树的空间复杂度都是 O(nlogn)。

之后我们要查询区间第 k 小的数字，我们先在根节点中确定对应的区间，再求出它对应的左子树的区间和右子树的区间。若 k 小于等于对应的左子树的区间长度，则第 k 大树在左子树中，否则在右子树中，之后向下递归直到叶节点即可。

但根据题目中的要求，我们不仅需要求出中位数，还要求出最小距离和。而要求出最小距离和，我们只要把范围内的数字分为小于中位数和大于中位数的两部分分别求和就可以了。考虑在划分树中查找的过程，若到某一节点时，第 k 小的数在其左子树内，我们要向左子树走，此时右子树中所有的数都大于这个第 k 小的数，我们对右子数对应范围内的数求和即可，求和使用部分和的算法，第 k 小数在右边时同理。

最后总的时间复杂度为 O(nlogn+qlogn)。

四、题目来源

我原本是看到了一道求最小曼哈顿距离和的题目，然后再变为这道题的。原题目用暴力的方法就能解决，而我把数据范围进行了扩大并且在外面又套了个“最小差异”的外壳。

原题目为 pku3269，来自 USACO 2007 Feb Gold，名为 Building A New Barn。

链接：http://poj.org/problem?id=3269