Code Jam 2015 - Round 3

Analysis: Smoothing Window

One approach towards a solution is to construct the temperature sequence that begins with K-1 zeroes, followed by a single pass through all the N-K+1 smoothing window sums to fill in the rest of the sequence. For example, running this construction for the sum array â€œ0 12 0 12 0â€ with K = 3 (from the third sample case) produces the sequence â€œ0 0 0 12 -12 12 0â€, where the difference between the maximum and minimum temperatures is 24. To reach a sequence that minimizes this difference, we tweak some values in the sequence as explained below.

Letâ€™s define GROUP(i) where (0 â‰¤ i < K) as the group containing the (i % K)-th temperatures in the sequence. The sequence â€œ0 **0** 0 12 **-12** 12 0â€ above is grouped as follows:

1. â€œ0 12 0â€ from the first, fourth, and seventh temperatures
2. â€œ**0 -12**â€ from the second and fifth temperatures
3. â€œ0 12â€ from the third and sixth temperatures

Changing the i-th temperature by Z degrees means that we need to compensate for this addition with other temperatures in the sequence. One way is by adding Z to the other temperatures in the same group, and subtracting Z from all the temperatures in another group. For example, to increase the first temperature reading by 3 degrees, we also increase the fourth and seventh temperature readings by 3, and reduce the temperatures in the second (or third) group by 3. The resulting sequence is â€œ3 -3 0 15 -15 12 3â€. Notice that its sum array stays the same: â€œ0 12 0 12 0â€.

Now letâ€™s define:

- lo(i) as the minimum element in GROUP(i)
- hi(i) as the maximum element in GROUP(i)
- interval(i) as a range [lo(i), hi(i)]
- SHIFT(i, y) as an operation that adds y to each member of GROUP(i).

We can remodel the original problem as follows:

> You are given K intervals, where the i-th interval spans from lo(i) to hi(i). You can perform any number of interval adjustments by choosing i, j, and y (0 â‰¤ i, j < K), shifting the i-th interval by y (i.e., SHIFT(i, y)), and shifting the j-th interval by -y (i.e., SHIFT(i, -y)). Find a sequence of adjustments (where order doesnâ€™t matter) that minimizes the covering range of all intervals, where the covering range of the intervals is max{hi(i)} - min{lo(j)} for all 0 â‰¤ i, j < K. The minimum covering range is equivalent to the smallest possible difference between the minimum and maximum temperatures in the original problem.

Another important insight is that the shifting can be done independently and can be aggregated. That is, we can accumulate all positive shifts and do it in bulk, similarly with the negative shifts. Therefore, we can â€œnormalizeâ€ all intervals by shifting them such that the lower bound of all intervals becomes 0. Denote Q to be the total sum of shifts, which is the sum of all lo(i) where 0 â‰¤ i < K. For example, if we have two intervals [-10, -8] and [333, 777], after normalization they become [0, 2] and [0, 444] with Q = -10 + 333 = 323. Finally, we have to shift back the normalized intervals by Q degrees, but with the flexibility to distribute the shifts such that the covering range of all intervals is minimized.

Notice that if we increment (or decrement) all the intervals by one, the relative positioning of each interval to one another won't change and thus it does not change the covering range of the intervals. Using this insight, we can reduce Q to Q % K by distributing the excess shifts Q evenly among the intervals without affecting the final answer. When Q is negative we can keep adding K to Q until it is non-negative.

Suppose we have three normalized intervals [0, 4], [0, 9], and [0, 7], and the total sum of shifts Q is 40. We can distribute back 39 shifts evenly to all three intervals where each interval is shifted by 13, ending with [13, 17], [13, 22], and [13, 20]. Then, we can shift all of them down back to [0, 4], [0, 9], [0, 7] without changing their relative positioning and and we are left with Q = 1.

Finally, we must distribute back the remaining shifts to the intervals and minimize the covering range of all intervals.

Suppose that L is the size of the largest interval. If Q = 0, then there is no excess shift to distribute, and the minimum covering range is L. Otherwise, we can assign L - (hi(i) - lo(i)) excess shifts to interval i without changing the answer. Using the previous example, the current minimum difference is L = 9. Without changing the minimum covering range, we can assign 1 excess shift to interval [0, 4], and get [1, 5].

If there are still some excess shifts remaining, it means we cannot distribute any more excess shifts without increasing the minimum covering range. Increasing the minimum covering range by 1 will make room for distributing K more excess shifts, which is enough for the remaining Q (since Q = sum of shifts % K, and thus it is less than K).

The complexity of this solution is O(N).