17CS10056
Ayush Tiwari

Tutorial 1

Problem Statement

$A[1..m]$ and $B[1..n]$ are two 1D arrays containing m and n integers respectively, where $m \leq n$. We need to construct a sub-array $Z[1..m]$ of B such that $\sum_{i=1}^{m} \big|A[i] - Z[i]\big|$ is minimized.

Recurrences

$X(i,j)$ is the minimum value of $\sum_{k=1}^{i} \big|A[i] - Z[i]\big|$ for the for the arrays $A_{\mathrm{1..i}}$ and $B_{\mathrm{1..j}}$

$$X(i,j) = \left\{ \begin{array}{lr} 0, & \text{for } i = 0 \\ \infty & \text{for } j < i \\ min(|B_{\mathrm{j}} - A_{\mathrm{i}}| + X(i-1, j-1), X(i, j-1)) & \text{for } i \leq j \leq n - m + i \end{array} \right\}$$

Algorithm

In a brute force approach to solving the Best Subarray Problem, we would enumerate all the subsequences of the array B and loop over all of these keeping track of the sequence that results in the minimum sum of differences.
It can easily seen that this problem has an optimal substructure property. For example, two contending subsequences might only differ by one element, but we end up looping over all the elements to calculate the sum of difference.

Optimal Substructure of the Best Sub Array Problem

Let $A = \langle a_1, a_2, a_3, ...a_{\mathrm{m}}\rangle$, $B = \langle b_1, b_2, b_3, ...b_{\mathrm{n}}\rangle$ and $m \leq n$
Let $Z = \langle z_1, z_2, z_3, ...z_{\mathrm{m}}\rangle$ be the Best Subarray of A and B.

1. If $z_{\mathrm{m}} = b_{\mathrm{n}}$ then $Z_{\mathrm{1..m-1}}$ must the Best Subarray of $A_{\mathrm{1..m-1}}$ and $B_{\mathrm{1..n-1}}$
2. If $z_{\mathrm{m}} \neq b_{\mathrm{n}}$ then Z is the Best Subarray of A and $B_{\mathrm{1..n-1}}$

Proof

1. If $z_{\mathrm{k}} = b_{\mathrm{n}}$, $\big|a_n - z_n\big|$ must have contributed to the sum, so we are left with the elements to the left of b_n and to the left of a_n. If $Z_{\mathrm{1..m-1}}$ is not the Best Subarray of $A_{\mathrm{1..m-1}}$ and $B_{\mathrm{1..n-1}}$ then we could find $Z'_{\mathrm{1..m-1}}$ for which $\sum_{i=1}^{m-1} \big|A[i] - Z[i]\big|$ is lesser and add it to $\big|a_n - z_n\big|$, thereby finding a better solution, which is a contradiction since Z is the Best Subarray.
2. This is obvious because if we cannot include b_{n} in Z then Z must be the Best Subarray of A and $B_{\mathrm{1..n-1}}$.

Psuedocodes

Data: Array B of length n and Array A of length $m \leq n$
Result: $sum = \sum_{i=1}^{m} \left|A[i] - Z[i]\right|$

```
begin
    X, Y ⟵ Matrix of dimension m + 1 × n + 1
    i, j ⟵ 0
    for i ≤ m do
        for j ≤ n − m + i do
            if i = 0 then
                X[i, j] = 0
            else if j < i then
                X[i, j] = ∞
            else
                if |B_j − A_i| + X[i − 1, j − 1] < X[i, j − 1] then
                    X[i, j] ⟵ |B_j − A_i| + X[i − 1, j − 1]
                    Y[i, j] ⟵ '↖'
                else
                    X[i, j] ⟵ X[i, j − 1]
                    Y[i, j] ⟵ '←'
                end
            end
        end
    end
end
```

Algorithm 1: Algorithm to calculate the minimum sum.

Data: Array Y calculated in Algorithm 1 and Array B of length n
Result: A vector containing the elements of Z and corresponding index in B

```
begin
    i ⟵ m, j ⟵ n
    v ⟵ emptyvector
    for i > 0 and j > 0 do
        if Y[i, j] = ↖ then
            v.pushfront((B[j], j))
            i = i − 1
            j = j − 1
        else
            j = j − 1
        end
    end
end
```

Algorithm 2: Algorithm to find the values of Z.

Demonstration 1

$A = \begin{bmatrix} 2 & 8 & 4 \end{bmatrix}$

$B = \begin{bmatrix} 5 & 9 & -6 & 1 \end{bmatrix}$

$$X^{1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ \infty & 0 & 0 & 0 & 0 \\ \infty & \infty & 0 & 0 & 0 \\ \infty & \infty & \infty & 0 & 0 \end{bmatrix}$$

Algorithm Steps

Iteration 1 : $i = 1, j = 1$

$X[0,0] + \big|A[1] - B[1]\big| = 3$

$X[1,0] = \infty$

Since $3 < \infty$, $X[0,0] = 3$ and $Y[0,0] = '\nwarrow'$

$$X = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ \infty & 3 & 0 & 0 & 0 \\ \infty & \infty & 0 & 0 & 0 \\ \infty & \infty & \infty & 0 & 0 \end{bmatrix}$$

Iteration 2 : $i = 1, j = 2$

$X[0,1] + \big|A[1] - B[2]\big| = 7$

$X[1,1] = 3$

Since $7 > 3$, $X[1,2] = 3$ and $Y[1,2] = '\leftarrow'$

$$X = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ \infty & 3 & 3 & 0 & 0 \\ \infty & \infty & 0 & 0 & 0 \\ \infty & \infty & \infty & 0 & 0 \end{bmatrix}$$

Iteration 3 : $i = 1, j = 3$

$X[0,2] + \big|A[1] - B[3]\big| = 8$

$X[1,2] = 3$

Since $8 > 3$, $X[1,3] = 3$ and $Y[1,3] = '\leftarrow'$

$$X = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ \infty & 3 & 3 & 3 & 0 \\ \infty & \infty & 0 & 0 & 0 \\ \infty & \infty & \infty & 0 & 0 \end{bmatrix}$$

Iteration 4 : $i = 1, j = 4$

$X[0,3] + \big|A[1] - B[4]\big| = 1$

$X[1,3] = 3$

Since $1 < 3$, $X[1,4] = 1$ and $Y[1,4] = '\leftarrow'$

$$X = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ \infty & 3 & 3 & 3 & 1 \\ \infty & \infty & 0 & 0 & 0 \\ \infty & \infty & \infty & 0 & 0 \end{bmatrix}$$

[1] By initializing X this way can skip the parts of the iteration when $i > j$ and $i = 0$

Iteration 5 : $i = 2, j = 2$

$X[1,1] + \big|A[2] - B[2]\big| = 4$

$X[2,1] = \infty$

Since $4 < \infty$, $X[2,2] = 4$ and $Y[2,2] = '\nwarrow'$

$$X = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ \infty & 3 & 3 & 3 & 1 \\ \infty & \infty & 4 & 0 & 0 \\ \infty & \infty & \infty & 0 & 0 \end{bmatrix}$$

Iteration 6 : $i = 2, j = 3$

$X[1,2] + \big|A[2] - B[3]\big| = 5$

$X[2,2] = 4$

Since $4 < 5, X[2,3] = 4$ and $Y[2,3] = '\leftarrow'$

$$X = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ \infty & 3 & 3 & 3 & 1 \\ \infty & \infty & 4 & 4 & 0 \\ \infty & \infty & \infty & 0 & 0 \end{bmatrix}$$

Iteration 7 : $i = 2, j = 4$

$X[1,3] + \big|A[2] - B[4]\big| = 10$

$X[2,3] = 4$

Since $4 < 10, X[2,3] = 4$ and $Y[2,3] = '\leftarrow'$

$$X = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ \infty & 3 & 3 & 3 & 1 \\ \infty & \infty & 4 & 4 & 4 \\ \infty & \infty & \infty & 0 & 0 \end{bmatrix}$$

Iteration 8 : $i = 3, j = 3$

$X[2,2] + \big|A[3] - B[3]\big| = 14$

$X[3,2] = \infty$

Since $14 < \infty, X[2,3] = 14$ and $Y[2,3] = '\nwarrow'$

$$X = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ \infty & 3 & 3 & 3 & 1 \\ \infty & \infty & 4 & 4 & 4 \\ \infty & \infty & \infty & 14 & 0 \end{bmatrix}$$

Iteration 9 : $i = 3, j = 4$

$X[2,3] + \big|A[3] - B[4]\big| = 7$

$X[3,3] = 14$

Since $7 < 14, X[3,4] = 7$ and $Y[3,4] = '\nwarrow'$

$$X = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ \infty & 3 & 3 & 3 & 1 \\ \infty & \infty & 4 & 4 & 4 \\ \infty & \infty & \infty & 14 & 7 \end{bmatrix}$$

Therefore the minimum value of $\sum_{i=1}^{3} \big|A[i] - Z[i]\big|$ *is* **7**.

Manual Check

1. $A = \begin{bmatrix} 2 & 8 & 4 \end{bmatrix}$
   $C = \begin{bmatrix} 5 & 9 & -6 \end{bmatrix}$
   $Sum = 14$

2. $A = \begin{bmatrix} 2 & 8 & 4 \end{bmatrix}$
   $C = \begin{bmatrix} 5 & 9 & 1 \end{bmatrix}$
   $Sum = 7$

3. $A = \begin{bmatrix} 2 & 8 & 4 \end{bmatrix}$
   $C = \begin{bmatrix} 5 & -6 & 7 \end{bmatrix}$
   $Sum = 20$

4. $A = \begin{bmatrix} 2 & 8 & 4 \end{bmatrix}$
   $C = \begin{bmatrix} 9 & -6 & 1 \end{bmatrix}$
   $Sum = 24$

Therefore the minimum sum is 7

Deducing the values of Z

A byproduct of the above algorithmic steps is the matrix Y

$$Y = \begin{bmatrix} - & - & - & - & - \\ - & \nwarrow & \leftarrow & \leftarrow & \nwarrow \\ - & - & \nwarrow & \leftarrow & \leftarrow \\ - & - & - & \nwarrow & \nwarrow \end{bmatrix}$$

Algorithmic Steps

1. At $Y[3,4]$, $B[4]$ is pushed front and $i \leftarrow 2, j \leftarrow 3$.
2. At $Y[2,3]$, $i \leftarrow 2, j \leftarrow 2$.
3. At $Y[2,2]$, $B[2]$ is pushed front and $i \leftarrow 1, j \leftarrow 1$.
4. At $Y[1,1]$, $B[1]$ is pushed front and $i \leftarrow 0, j \leftarrow 0$.

Demonstration 2

$B = \begin{bmatrix} 6 & 2 & 1 & 9 & 4 & 5 & -1 \end{bmatrix}$
$A = \begin{bmatrix} -2 & 6 & 3 & 8 \end{bmatrix}$

$$X^2 = \begin{bmatrix} 0 & 0 & 0 & 0 & - & - & - & - \\ \infty & 0 & 0 & 0 & 0 & - & - & - \\ \infty & \infty & 0 & 0 & 0 & 0 & - & - \\ \infty & \infty & \infty & 0 & 0 & 0 & 0 & - \\ \infty & \infty & \infty & \infty & 0 & 0 & 0 & 0 \end{bmatrix}$$

Results of the algorithm

$$X = \begin{bmatrix} 0 & 0 & 0 & 0 & - & - & - & - \\ \infty & 8 & 4 & 3 & 3 & - & - & - \\ \infty & \infty & 8 & 8 & 6 & 5 & - & - \\ \infty & \infty & \infty & 10 & 10 & 7 & 7 & - \\ \infty & \infty & \infty & \infty & 11 & 11 & 10 & 10 \end{bmatrix}$$

$$Y = \begin{bmatrix} - & - & - & - & - & - & - & - \\ - & \nwarrow & \nwarrow & \nwarrow & \leftarrow & - & - & - \\ - & - & \nwarrow & \leftarrow & \nwarrow & \nwarrow & - & - \\ - & - & - & \nwarrow & \leftarrow & \nwarrow & \leftarrow & - \\ - & - & - & - & \nwarrow & \leftarrow & \nwarrow & \leftarrow \end{bmatrix}$$

By following the same procedure in previous demonstration we find that the Best Subarray is

$Z = \begin{bmatrix} 1 & 9 & 4 & 5 \end{bmatrix}$

This can be easily verified by inspection.

[2]By initializing X this way can skip the parts of the iteration when $i > j$ and $i = 0$

Time and Space Complexities

1. **Algorithm 1** : The time complexity is $O(n \times m)$ and the sapce complexity is $O(n \times m)$.

   Justification

   In Algorithm 1, the outer for loop runs for m times and the inner for loop runs n times. Therefore the time complexity is $\Theta(n \times m)$.
   The X and Y matrix both occupy $m+1 \times n+1$ space, thus making the space complexity $\Theta(m \times n)$.

2. **Algorithm 2** : The time complexity is $O(n)$ and the sapce complexity is $O(m \times n)$.

   Justification

   In Algorithm 2, the loop will run for a maximum of n times, Therefore the time complexity is $\Theta(n)$. The Y matrix occupies $m + 1 \times n + 1$ space, thus making the space complexity $\Theta(m \times n)$.

Therefore the overall time complexity $= \Theta(m \times n)$ and the overall space complexity is $\Theta(m \times n)$.