思考题 7

孟妍廷 2015202009

2017 年 11 月 7 日

依题意可得，最优直方图问题符合最优子结构性质，适合用动态规划解决问题.
设 $SSE^*[i,k]$ 表示将排序后的数组中第 1 个到第 i 个元素放入 k 个桶中的最小二次求和误差.
AVG[i,j] 表示排序后的第 i 到第 j 个元素在一个桶中的均值.
算法如下
首先对原始数组 a[1...n] 由小到大排序，得到 order[1...n] 记录下标.

$COUNT(order, SSE*, AVG, PP, i, k)$
  $if\ k == 1$
    $cost = SSE*[j, 1]$
    $for\ j = 1\ to\ i-1$
      $cost = \infty$
      $temp = COUNT(order, SSE*, AVG, PP, j, k-1) + PP[i] - PP[j+1] - (i-j) \times AVG[j,i]^2$
      $(cost > temp)?cost = temp : cost = cost$
  $return\ cost$

$SOLUTION(order, n)$
  $let\ AVG[1...n, 1...n], SSE^*[1...n, 1...B], P[0...n], PP[0...n]\ be\ new\ array.$
  $P[0] = 0, PP[0] = 0$
  $for\ i = 1\ to\ n$
    $for\ j = 1\ to\ i$
      $P[i] += a[order[i]]$
      $PP[i] += a[order[i]]^2$
  $for\ i = 1\ to\ n$
    $for\ j = i\ to\ n$
      $AVG[i,j] = \frac{P[j]-P[i-1]}{j-i+1}$//计算出所有平均数
      $SSE^*[i,j] = \infty$
  $for\ i = 1\ to\ n$
    $SSE^*[i,1] = PP[i] - i \times AVG[1,i]^2$//所有数放入一个桶中的情况的误差可以直接算
  $for\ k = 1\ to\ B$
    $res = \infty$
    $if\ res > COUNT(order, SSE*, AVG, PP, n, k)$
      $res = COUNT(order, SSE*, AVG, PP, n, k)$
      $key = k$
  $return\ key\ res$

时间复杂度为 $O(B \times B \times n)$