

思考题 4: 最优直方图 (V-Optimal Histogram)

直方图的误差有很多定义, 其中最常用的定义是二次求和误差(sum squared error), 即每个数与其估计之差的平方和: $\sum_{i=1}^k \sum_{l=j_i}^{j_{i+1}-1} (\mu_i - a_l)^2$ 。使得二次求和误差最小的 k 个桶的划分,

就是最优直方图(V-Optimal histogram)。试利用动态规划, 设计算法求解最优直方图。

阅读: <http://infolab.stanford.edu/~datar/courses/cs361a/papers/vopt.pdf>