

アルゴリズム及び演習 第 8 回演習問題*

小野 孝男†

2007 年 6 月 11 日

以下の問いに答えよ:

1. n 個のデータ x_1, x_2, \dots, x_n に対し, 次の関数 $f(x)$ を考える:

$$f(x) = \sum_{i=1}^n |x - x_i|.$$

この関数 $f(x)$ は x が x_1, \dots, x_n の中央値のときに最小となることを示せ.

2. ここでは重み付き中央値を考える. n 個のデータの集合 $X = \{x_1, x_2, \dots, x_n\}$ と各 x_i に対する重み w_1, w_2, \dots, w_n が与えられる (但し $\sum_{i=1}^n w_i = 1$ と仮定する). このとき,

$$\sum_{i: x_i \leq x_{wm}} w_i \geq 1/2,$$
$$\sum_{i: x_i \geq x_{wm}} w_i \geq 1/2$$

を満たす $x_{wm} \in X$ を X の重み付き中央値と呼ぶ. この重み付き中央値を $O(n)$ 時間で求めるアルゴリズムを与えよ (ヒント: 中央値と同じようなアルゴリズムを考える).

3. 0 以上 $n^2 - 1$ 以下の n 個の整数が与えられたときに, これらを $O(n)$ 時間でソートする方法を示せ.
4. (レポート課題) n 個のデータの集合 $X = \{x_1, x_2, \dots, x_n\}$ に対し, k 分位点 $q_1, q_2, \dots, q_{k-1} \in X$ を次のように定義する: 各 i に対し, X において q_i 以下のデータは $\lfloor in/k \rfloor$ 個存在する. 例えば, $k = 4$ とすると q_1 以下のデータが $\lfloor n/4 \rfloor$ 個, q_2 以下のデータが $\lfloor n/2 \rfloor$ 個, q_3 以下のデータが $\lfloor 3n/4 \rfloor$ 個であるような q_1, q_2, q_3 を求めたい*¹. X と整数 k が与えられたときに, 全ての k 分位点を $O(n \log k)$ で求めるアルゴリズムを示せ.

* これまでの解答などは <http://www.al.cm.is.nagoya-u.ac.jp/~takao/lecture/> にあります.

† ono@is.nagoya-u.ac.jp

*¹ $k = 4$ の場合, q_1 を第 1 四分位点, q_3 を第 3 四分位点と呼ぶ. q_2 は中央値となる.