Median Sensitivity Proofs

May 21, 2020

Definition 1. *The sample median of a database $X = (x_1, \dots, x_n)$ is given by*

$$f(X) = \frac{\tilde{x}_l + \tilde{x}_u}{2}$$

where $l = \lfloor \frac{n+1}{2} \rfloor$ and $u = \lceil \frac{n+1}{2} \rceil$ and $\tilde{X} = (\tilde{x}_1, \dots, \tilde{x}_n)$ is the sorted version of X.

Note that this method of producing a DP median is not usable as the noise is scaled to the full range of the dataset.

1 Neighboring Definition: Change One

1.1 ℓ_1-sensitivity

Theorem 1. *Let the database X have elements bounded above by M and below by m. Then the ℓ_1-global sensitivity in the change-one model of the median is*

$$\Delta f(\cdot) = \begin{cases} \frac{M-m}{2}, & \text{if } n \equiv 0 \pmod 2 \\ M - m & \text{if } n \equiv 1 \pmod 2 \end{cases}$$

Proof. First consider the case where n is such that $n \equiv 0 \pmod 2$. Then our worst-case scenario is that exactly half of the data elements in X are m and the other half are M. In this case, $f(X) = \frac{m+M}{2}$.

Now consider an X' which is identical to X but with one additional element with value M.[1] Then we have that $f(X') = M$. Because we are looking at our worst-case pairing of X, X', we know that

$$\Delta f(\cdot) = \max_{X,X'} |f(X') - f(X)| = \left|M - \frac{m+M}{2}\right| = \frac{M-m}{2}.$$

Now consider the case where n is such that $n \equiv 1 \pmod 2$. Then our worst-case scenario is that $\frac{n-1}{2}$ of the elements of X are m, $\frac{n-1}{2}$ of the elements of X are M, and the remaining

[1] The result holds if it is switched from M to m.

element of X is m (WLOG). In this setting, $f(X) = m$. Let X' be identical to x but with one element switched from m to M. Then $f(X') = M$ and we have that

$$\Delta f(\cdot) = \max_{X,X'} |f(X') - f(X)| = |M - m|.$$

$\square$

1.2 ℓ_2-sensitivity

Theorem 2. *Let the database X have elements bounded above by M and below by m. Then the global ℓ_2-sensitivity in the change-one model of the median is*

$$\Delta f(\cdot) = \begin{cases} \frac{M-m}{2}, & \text{if } n \equiv 0 \pmod 2 \\ M - m & \text{if } n \equiv 1 \pmod 2 \end{cases}$$

Proof. The logic follows exactly from the proof of the ℓ_1 sensitivity, just with the norm in the last line of each statement switched from 1 to 2. $\square$

2 Neighboring Definition: Add/Drop One

2.1 ℓ_1-sensitivity

Theorem 3. *Let the database X have elements bounded above by M and below by m. Then the global ℓ_1-sensitivity in the add/drop-one model of the median is*

$$\Delta f(\cdot) = \frac{M-m}{2}.$$

Proof. First consider the case where n is such that $n \equiv 0 \pmod 2$. Then our worst-case scenario is that exactly half of the data elements in X are m and the other half are M. In this case, $f(X) = \frac{m+M}{2}$.

Now consider an X' which is identical to X but with a single additional element M.Then $f(X') = M$. Note that if instead X' was identical to X with a single element *removed*, the worst case is that a single m is removed and again $f(X') = M$.[2] In either case,

$$\left|f(X') - f(X)\right| = \left|M - \frac{M-m}{2}\right| = \frac{M-m}{2}.$$

Now consider the case where n is such that $n \equiv 1 \pmod 2$. Then our worst-case scenario is that $\frac{n-1}{2}$ of the elements of X are m, and $\frac{n+1}{2}$ of the elements of X are M (WLOG). In this setting, $f(X) = M$. Let X' be identical to x but with one element m added. Then $f(X') = \frac{M-m}{2}$. Note that if instead X' was identical to X with a single element removed, the worst case is that a single M is removed and again $f(X') = \frac{M-m}{2}$ (WLOG). In either case,

$$\left|f(X) - f(X')\right| = \left|M - \frac{M-m}{2}\right| = \frac{M-m}{2}.$$

So, in general,

$$\Delta f(\cdot) = \max_{X,X'} \left|f(X') - f(X)\right| = \frac{M-m}{2}.$$

$\square$

[2]The results holds if the point added/removed is switched from M to m.

2.2 ℓ_2-sensitivity

Theorem 4. *Let the database X have elements bounded above by M and below by m. Then the global ℓ_2-sensitivity in the add/drop-one model of the median is*

$$\Delta f(\cdot) = \frac{M - m}{2}.$$

Proof. The logic follows exactly from the proof of the ℓ_1 sensitivity, just with the norm switched from 1 to 2.

□