One-dimensional $l1$ cost function minimization

Proposition: Let $x_1, x_2, \dots, x_n \in R$, with $x_1 < x_2 < \cdots < x_n$. Consider the quantity

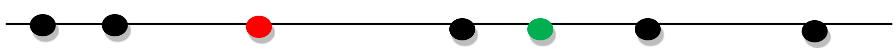
$$A = \sum_{i=1}^{n} |x_i - \mu|.$$

Then, A is minimized when μ is chosen as the median of $x_1, x_2, \dots, x_n$, i.e.

$$\mu = med(x_1, x_2, \dots, x_n).$$

Proof:

(a) Let n be **odd** and $x_q \equiv \mu = med(x_1, x_2, \dots, x_n)$. Consider a number $k > x_q$ so that $x_r \leq k < x_{r+1}$, with $r > q$ (see the following figure).



$$x_1 < x_2 < \cdots < x_q < x_{q+1} \dots < x_r < k < \quad x_{r+1} < \cdots < x_n.$$

Let $k = x_q + \lambda$, $(\lambda > 0)$.

Also, let

$$A_1 = \sum_{i=1}^{n} |x_i - x_q|,$$

and

$$A_2 = \sum_{i=1}^{n} |x_i - k|.$$

We consider the following three cases

(i) $x_i \leq x_q$: In this case it is $|x_i - x_q| = x_q - x_i$ and $|x_i - k| = k - x_i = x_q - x_i + \lambda$

(ii) $x_q < x_i \leq x_r (\leq k)$: In this case it is $|x_i - x_q| = x_i - x_q$ and $|x_i - k| = k - x_i = x_q - x_i + \lambda$

(iii) $x_r (\leq k) < x_i$: In this case it is $|x_i - x_q| = x_i - x_q$ and $|x_i - k| = x_i - k = x_i - x_q - \lambda$

It is

$$A_1 = \sum_{i=1}^{q} |x_i - x_q| + \sum_{i=q+1}^{r} |x_i - x_q| + \sum_{i=r+1}^{n} |x_i - x_q|$$
$$= \sum_{i=1}^{q} (x_q - x_i) + \sum_{i=q+1}^{r} (x_i - x_q) + \sum_{i=r+1}^{n} (x_i - x_q)$$

and

$$A_2 = \sum_{i=1}^{q} |x_i - k| + \sum_{i=q+1}^{r} |x_i - k| + \sum_{i=r+1}^{n} |x_i - k|$$
$$= \sum_{i=1}^{q} (x_q - x_i + \lambda) + \sum_{i=q+1}^{r} (x_q - x_i + \lambda) + \sum_{i=r+1}^{n} (x_i - x_q - \lambda)$$
$$= \sum_{i=1}^{q} (x_q - x_i) + \sum_{i=1}^{q} \lambda + \sum_{i=q+1}^{r} (x_q - x_i) + \sum_{i=q+1}^{r} \lambda + \sum_{i=r+1}^{n} (x_i - x_q)$$
$$- \sum_{i=r+1}^{n} \lambda$$

Then, it is

$$\Lambda = A_2 - A_1 = \sum_{i=1}^{q} \lambda + \sum_{i=q+1}^{r} \lambda - \sum_{i=r+1}^{n} \lambda + \sum_{i=q+1}^{r} (x_q - x_i) - \sum_{i=q+1}^{r} (x_i - x_q)$$
$$= \sum_{i=1}^{r} \lambda - \sum_{i=r+1}^{n} \lambda + 2 \sum_{i=q+1}^{r} (x_q - x_i) = (2r - n)\lambda - 2 \sum_{i=q+1}^{r} (x_i - x_q)$$

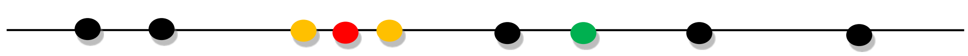
For $i = q + 1, \dots, r$, it is $x_i - x_q \le x_r - x_q \le k - x_q = \lambda$.

Therefore (taking also into account that $q = \frac{n+1}{2}$)

$$\Lambda = A_2 - A_1 \ge (2r - n)\lambda - 2 \sum_{i=q+1}^{r} \lambda = (2r - n)\lambda - 2(r - q)\lambda = (2q - n)\lambda = \left(2\frac{n+1}{2} - n\right)\lambda = \lambda > 0$$

The case where $k > x_q$ is treated similarly.

(b) Let n be **even** and $q = \frac{n}{2}$. Then, the median is $\frac{x_q + x_{q+1}}{2}$.



$$x_1 < x_2 < \cdots < x_q < x_{q+1} \dots < x_r < k < \quad x_{r+1} < \cdots < x_n.$$

We proceed as follows:

(i) for any $k < x_q$ we prove that

$$\sum_{i=1}^{n} |x_i - x_q| < \sum_{i=1}^{n} |x_i - k|$$

(ii) for any $k > x_{q+1}$ we prove that

$$\sum_{i=1}^{n} |x_i - x_{q+1}| < \sum_{i=1}^{n} |x_i - k|$$

(iii) for any $k, \mu \in [x_q, x_{q+1}]$ we prove that

$$\sum_{i=1}^{n} |x_i - k| = \sum_{i=1}^{n} |x_i - \mu|$$

The (i) and (ii) can be proved using the rationale adopted in the case where n is odd.

For the (iii) case, assuming that $k < \mu$, we have

$$B_1 = \sum_{i=1}^{n} |x_i - k| = \sum_{i=1}^{q} |x_i - k| + \sum_{i=q+1}^{n} |x_i - k| = \sum_{i=1}^{q} (k - x_i) + \sum_{i=q+1}^{n} (x_i - k)$$

and

$$B_2 = \sum_{i=1}^{n} |x_i - \mu| = \sum_{i=1}^{q} |x_i - \mu| + \sum_{i=q+1}^{n} |x_i - \mu| = \sum_{i=1}^{q} (\mu - x_i) + \sum_{i=q+1}^{n} (x_i - \mu)$$

Taking the difference between B_1 and B_2, we have

$$B_1 - B_2 = \sum_{i=1}^{q} [(k - x_i) - (\mu - x_i)] + \sum_{i=q+1}^{n} [(x_i - k) - (x_i - \mu)] = \sum_{i=1}^{q} (k - \mu) + \sum_{i=q+1}^{n} (\mu - k)$$
$$= q(k - \mu) + (n - q)(\mu - k) = \frac{n}{2}(k - \mu) + \left(n - \frac{n}{2}\right)(\mu - k) = 0$$

Therefore, all $k \in [x_q, x_{q+1}]$ minimize the quantity

$$A = \sum_{i=1}^{n} |x_i - k|.$$