Mean minimizes L^2 norm

The minimum of

$$f(c) = \sum_{i=1}^{n} (x_i - c)^2$$

happens at $c = \frac{1}{n} \sum_{i=1}^{n} x_i$ (the average of the points).

The idea is to **multiply everything out and complete the square**. I.e., write this function in the form $f(c) = a((c - q)^2 + p)$. Then the minimum will occur at q.

Simple motivating example.

Let $x_1 = 3, x_2 = 5$. Let's write $(c - x_i)^2$ instead of $(x_i - c)^2$ (this has no effect because of the squaring). So we want to minimize

$$(c - 3)^2 + (c - 5)^2.$$

Multiply everything out and obtain

$$c^2 - 6c + 9 + c^5 - 10c + 25 = 2c^2 - 16c + 34.$$

Factor out a 2, and add a clever form of 0:

$$2(c^2 - 8c + 17) = 2((c^2 - 8c + 16) - 16 + 17).$$

Of course, $c^2 - 8c + 16 = (c - 4)^2$, so this is simply

$$2((c - 4)^2 + 1).$$

The minimum clearly occurs at $c = 4$.

More general proof.

Write

$$f(c) = \sum_{i=1}^{n} (c - x_i)^2 = \sum_{i=1}^{n} (c^2 - 2cx_i + x_i^2).$$

This is

$$= nc^2 - 2c \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} x_i^2$$

$$= n(c^2 - (\frac{2}{n} \sum_{i=1}^{n} x_i)c + \frac{1}{n} \sum_{i=1}^{n} x_i^2).$$

Now complete the square by adding a clever form of 0 : $(\frac{1}{n} \sum_{i=1}^{n} x_i)^2 - (\frac{1}{n} \sum_{i=1}^{n} x_i)^2$. Thus we obtain

$$n\left(\left(c^2 - \left(\frac{2}{n}\sum_{i=1}^{n} x_i\right)c + \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^2 + \left(\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^2\right)\right)\right).$$

The first three terms can be rewritten nicely:

$$n\left(\left(c - \frac{1}{n}\sum_{i=1}^{n} x_i\right)^2 + \left(\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^2\right)\right).$$

This clearly attains a minimum at $c = \frac{1}{n} \sum_{i=1}^{n} x_i$, which is the mean. So we can minimize this expression by computing the (arithmetic) mean in linear time.

One could also show that the constant $\frac{1}{n} \sum_{i=1}^{n} x_i^2 - \left(\frac{1}{n} \sum_{i=1}^{n} x_i\right)^2$ is ≥ 0. Simply apply Cauchy-Schwarz with vectors $\vec{u} = (x_1, \ldots, x_n), \vec{v} = (1/n, \ldots, 1/n)$. (Though this is unnecessary for this problem.)