On the Unreasonable Efficacy of the Mean in Minimizing the Fuel Expenditure of Crab Submarines

CrashAndSideburns

December 7, 2021

Abstract

In the plotting of courses for crab submarines, as well as in some novel techniques for escaping from angry whales, it is occasionally important to consider the problem of minimizing fuel expenditure. In this paper, we discuss the problem of minimizing the fuel expenditure required to bring a set of n crab submarines to a single point. This problem is of particular relevance to novel whale-escape techniques, as a large number of crab submarines are required at the same point to blast sufficiently large holes in the ocean floor tor these techniques to be viable. Finally, it was noticed by a group of participants in the 2021 Advent of Code challenge that the mean provides a very near estimate of optimal points for the minimization of fuel expenditure. but is not always optimal. The reason for this approach's unreasonable efficacy and ultimate inadequacy are also discussed.

1 On the Dynamics of Crab Submarines

Let us begin by considering a group of crab submarines with some arbitrary position, which we denote x_i for the ith crab submarine. A set of n crab submarines may then be denoted as the n-tuple $(x_1, \ldots, x_n)$. We must first consider the fuel expended by this set of crab submarines in moving to another point k. To do so, we consider an arbitrary crab submarine, x_i. It is well-known that the fuel expenditure of a crab submarine moving from x_i to k, which we denote $f_i(k)$, may be written as

$$
\begin{aligned}
f_i(k) &= 1 + 2 + 3 + \ldots + |k - x_i| \\
&= \sum_{n=0}^{|k-x_i|} n \\
&= \frac{|k - x_i| \left(|k - x_i| + 1\right)}{2} \\
&= \frac{(k - x_i)^2 + |k - x_i|}{2}
\end{aligned}
$$

where the closed form of the series is due to a result by Gauss. It must then be the case that the fuel expenditure of all n crab submarines may be obtained by summing the expenditures of each individual submarine, so the total fuel expenditure of the crab submarines, $f(k)$, is given by

$$
\begin{aligned}
f(k) &= \sum_{i=1}^{n} f_i(k) \\
&= \sum_{i=1}^{n} \frac{(k - x_i)^2 + |k - x_i|}{2}
\end{aligned}
$$

which is the amount of fuel expended by a set of n crab submarines with initial positions x_i in moving to a point k.

2 On the Minimization of Fuel Expenditure

We are now at a point where we have an expression for the fuel expenditure of a set of crab submarines, and we wish to find some point k to which the crab submarines may be moves which minimizes the total fuel expenditure. In other words, given $f(k)$ we wish to find some k which minimizes f. This is a very common problem in calculus, and it is known that the minimum of f, if one exists, must occur at a point where

$$\frac{\mathrm{d}f}{\mathrm{d}k} = 0$$

Given the definition of f found in part 1 of this paper, the derivative may be readily calculated, and we find

$$\frac{\mathrm{d}f}{\mathrm{d}k} = \sum_{i=1}^{n} \frac{2(k - x_i) + \operatorname{sgn}(x_i - k)}{2}$$

where $\operatorname{sgn}(x)$ is a function called the *sign function*, and is equal to 1 if $x > 0$, 0 if $x = 0$, and -1 if $x < 0$. We are looking for the value of k which minimizes f, which occurs when

$$\sum_{i=1}^{n} \frac{2(k - x_i) + \operatorname{sgn}(x_i - k)}{2} = 0$$

The usefulness of the above equation may not be immediately obvious, but by rearranging terms we may put it in the form

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} \left(k + \frac{\operatorname{sgn}(x_i - k)}{2} \right)$$

at which point k may be drawn out of the sum to yield

$$\sum_{i=1}^{n} x_i = nk + \sum_{i=1}^{n} \frac{\operatorname{sgn}(x_i - k)}{2}$$

which gives us an expression for k that is very nearly in a closed form, namely that

$$k = \frac{\sum_{i=1}^{n} x_i}{n} - \frac{1}{2} \frac{\sum_{i=1}^{n} \operatorname{sgn}(x_i - k)}{n}$$

It is now valuable to note that $\frac{\sum_{i=1}^{n} x_i}{n}$ is precisely the mean of the x_i, which we denote $\overline{x}$. This gives us the simplified form

$$k = \overline{x} - \frac{1}{2} \frac{\sum_{i=1}^{n} \operatorname{sgn}(x_i - k)}{n}$$

but unfortunately another term persists, and remains dependent on k. Eliminating this term is intractable, but it may be bounded. Namely, we observe that $\sum_{i=1}^{n} \operatorname{sgn}(x_i - k)$ is maximally n when $x_i > k$ for all i, and minimally $-n$, when $x_i < k$ for all i. As such, the second term is maximally equal to $\frac{1}{2}$ and minimally equal to $-\frac{1}{2}$. This means that while k may not be exactly calculated, it is bounded by

$$\overline{x} + \frac{1}{2} \geq k \geq \overline{x} - \frac{1}{2}$$

which provides a sufficiently tight bound for the numerical calculation of optimal crab submarine courses.