Optimal Search under Constraints

Martine Ceberio, Olga Kosheleva, and Vladik Kreinovich

Abstract In general, if we know the values a and b at which a continuous function has different signs – and the function is given as a black box – the fastest possible way to find the root x for which $f(x) = 0$ is by using bisection (also known as binary search). In some applications, however – e.g., in finding the optimal dose of a medicine – we sometimes cannot use this algorithm since, to avoid negative side effects, we can only try values which exceed the optimal dose by no more than some small value $\delta > 0$. In this paper, we show how to modify bisection to get an optimal algorithm for search under such constraint.

1 Where This Problem Came From

Need to select optimal dose of a medicine. This research started with a simple observation about how medical doctors decide on the dosage. For many chronic health conditions like high cholesterol, high blood pressure, etc., there are medicines that bring the corresponding numbers back to normal. An important question is how to select the correct dosage:

- on the one hand, if the dosage is too small, the medicine will not have the full desired effect;
- on the other hand, we do not want the dosage to be higher than needed: every medicine has negative side effects, side effects that increase with the increase in dosage, and we want to keep these side effects as small as possible.

In most such cases, there are general recommendations that provide a range of possible doses depending on the patient's age, weight, etc., but a specific dosage within

---

Martine Ceberio, Olga Kosheleva, and Vladik Kreinovich
University of Texas at El Paso, 500 W. University, El Paso, TX 79968, USA
e-mail: mceberio@utep.edu, olgak@utep.edu, vladik@utep.edu

this range has to be selected individually, based on how this patient's organism reacts to this medicine.

How the first doctor selected the dose. It so happened that two people having similar conditions ended up with the same daily dosage of 137 units of medicine, but interestingly, their doctors followed a different path to this value.

For the first patient, the doctor seems to have followed the usual bisection algorithm:

- this doctor started with the dose of 200 – and it worked,
- so, the doctor tried 100 – it did not work,
- the doctor tried 150 – it worked,
- the doctor tried 125 – it did not work,
- so, the doctor tried 137 – and it worked.

The doctor could have probably continued further, but the pharmacy already had trouble with maintaining the exact dose of 137, so this became the final arrangement.

This procedure indeed follows the usual bisection (= binary search) algorithm (see, e.g., [1]) – which is usually described as a way to solve the equation $f(x) = 0$ when we have an interval $[a,b]$ for which $f(a) < 0 < f(b)$. In our problem, $f(a)$ is the difference between the effect of the dose a and the desired effect:

- if the dose is not sufficient, this difference is negative, and
- if the dose is sufficient, this difference is non-negative (positive or 0).

In the bisection algorithm, at each iteration, we have a range $[\underline{x},\overline{x}]$ for which $f(\underline{x}) < 0$ and $f(\overline{x}) > 0$. In the beginning, we have $[\underline{x},\overline{x}] = [a,b]$. At each iteration, we take a midpoint $m = \frac{\underline{x}+\overline{x}}{2}$ and compute $f(m)$. Depending on the sign of $f(m)$, we make the following changes:

- if $f(m) < 0$, we replace $\underline{x}$ with m and thus, get a new interval $[m,\overline{x}]$;
- if $f(m) > 0$, we replace $\overline{x}$ with m and thus, get a new interval $[\underline{x},m]$.

In both cases, we decrease the width of the interval $[\underline{x},\overline{x}]$ by half. We stop when this width becomes smaller than some given value $\varepsilon > 0$; this value represents the accuracy with which we want to find the solution.

In the above example, based on the first experiment, we know that the desired dose is within the interval $[0,200]$. So:

- we try $m = 100$ and, after finding that $f(m) < 0$ (i.e., that the dose $m = 100$ is not sufficient), we come up with the narrower interval $[100,200]$;
- then, we try the new midpoint $m = 150$, and, based on the testing result, we come up with the narrower interval $[100,150]$;
- then, we try the new midpoint $m = 125$, and, based on the testing result, we come up with the narrower interval $[125,150]$;
- in the last step, we try the new midpoint $m = 137$ (strictly speaking, it should be 137.5, but, as we have mentioned, the pharmacy cannot provide such an accuracy); now we know that the desired value is within the narrower interval

$$[125,137].$$

Out of all possible values from the interval $[125, 137]$, the only value about which we know that this value is sufficient is the value 137, so this value has been prescribed to the first patient.

The second doctor selected the same dose differently. Interestingly, for the second patient, the process was completely different:

- the doctor started with 25 units;
- then – since this dose was not sufficient – the dose was increased to 50 units;
- then the dose was increased to 75, 100, 125 units, and, finally, to 150 units.

The 150 units dose turned out to be sufficient, so the doctor knew that the optimal dose is between 125 and 150. Thus, this doctor tried 137, and it worked.

Comment. Interestingly, in contrast to the first doctor, this doctor could not convince the pharmacy to produce a 137 units dose. So this doctor's prescription of this dose consists of taking 125 units and 150 units in turn.

Why the difference? Why did the two doctors use different procedures?

Clearly, the second doctor needed more steps – and longer time – to come up with the same optimal dose: this doctor used 7 steps (25, 50, 75, 100, 125, 150, 137) instead of only 5 steps used by the first doctor (200, 100, 150, 125, 137). Why did this doctor not use a faster bisection procedure?

At first glance, it may seem that the second doctor was not familiar with bisection – but clearly this doctor *was* familiar with it, since, after realizing that the optimal dose is within the interval $[125, 150]$, he/she checked the midpoint dose of 137.

The real explanation of why the second doctor did not use the faster procedure is that the second doctor was more cautious about possible side effects – probably, in this doctor's opinion, the second patient was vulnerable to possible side effects. Thus, this doctor decided not to increase the dose too much beyond the optimal value, so as to minimize possible side effects – while the first doctor, based on the overall health of the first patient, was less worried about possible side effects.

Natural general question. A natural next question is: under such restriction on possible tested values x, what is the optimal way to find the desired solution (i.e., to be more precise, the desired ε-approximation to the solution)?

It is known that if we do not have any constraints, then bisection is the optimal way to find the solution to the equation $f(x) = 0$; see, e.g., [1]. So, the question is – how to optimally modify bisection under such constraints?

2 Towards Formulating the Problem in Precise Terms

The larger the dose of the medicine, the larger the effect. There is a certain threshold x_0 after which the medicine has the full desired curing effect.

Every time we test a certain does x of the medicine of a patient:

- we either get the full desired effect, which would mean that $x_0 \le x$,

- or we do not yet get the full desired effect, which means that $x < x_0$.

We want to find the curing dose as soon as possible, i.e., after as few tests as possible.

If the only objective was to cure the disease, then, in principle, we could use any dose larger than or equal to x_0. However, the larger the dose, the larger the undesired side effects. So, we would like to prescribe a value which is as close to x_0 as possible. Of course, in real life, we can only maintain the dose with some accuracy $\varepsilon > 0$. So, we want to prescribe a value x_r which is ε-close to x_0, i.e., for which $x_0 \le x_r \le x_0 + \varepsilon$.

The only way to find the optimal dose x_r is to test different doses on a given patient. If, during this testing, we assign too large a dose, we may seriously harm the patient. So, it is desirable not to exceed x_0 too much when testing. Let us denote the largest allowed excess by δ. This means that we can only test values $x \le x_0 + \delta$.

Now, we can formulate the problem in precise terms.

3 Precise Formulation of the Problem and the Optimal Algorithm

Definition 1. *By* search under constraints*, we mean the following problem:*

- *Given:*
  - *rational numbers $\varepsilon > 0$ and $\delta > 0$, and*
  - *an algorithm c that, for some fixed (unknown) value $x_0 > 0$, given a rational number $x \in [0, x_0 + \delta]$, checks whether $x < x_0$ or $x \ge x_0$; this algorithm will be called a* checking *algorithm.*
- *Find: a real number x_r for which $x_0 \le x_r \le x_0 + \varepsilon$.*

Comment. We want to find the fastest possible algorithm for solving this problem. To gauge the speed of this algorithm, we will count the number of calls to the checking algorithm c.

Definition 2.

- *For every algorithm A for solving the search under constraints problem, let us denote the number of calls to the checking algorithm c corresponding to each instance $(\varepsilon, \delta, x_0)$ by $N_{A,\varepsilon,\delta}(x_0)$.*
- *We say that the algorithm A_0 for solving the search under constraint problem is* optimal *if for each ε and δ, the function $N_{A_0,\varepsilon,\delta}(x_0)$ is asymptotically optimal, i.e., that for every other algorithm A for solving the search under constraints problem, we have*

$$N_{A_0,\varepsilon,\delta}(x_0) \le N_{A,\varepsilon,\delta}(x_0) + \text{const}$$

for some constant depending on ε, δ, and A.

Proposition. *The following algorithm $\mathscr{A}$ is optimal:*

- *First, we apply the algorithm c to values δ, 2δ, ..., until we find a value i for which $i \cdot \delta < x_0 \le (i+1) \cdot \delta$.*
- *Then, we apply bisection process to the interval $[i \cdot \delta, (i+1) \cdot \delta]$ to find x_r:*
  - *In this process, at each moment of time, we have an interval $[\underline{x}, \overline{x}]$ for which*
  $$\underline{x} < x_0 \le \overline{x}.$$
  - *We start with $[\underline{x}, \overline{x}] = [i \cdot \delta, (i+1) \cdot \delta]$.*
  - *At each iteration step, we apply the checking algorithm c to the midpoint*
  $$m = \frac{\underline{x} + \overline{x}}{2}.$$
  - *If it turns out that $m < x_0$, we replace $[\underline{x}, \overline{x}]$ with $[m, \overline{x}]$.*
  - *If it turns out that $x_0 \le m$, we replace $[\underline{x}, \overline{x}]$ with $[\underline{x}, m]$.*
  - *In both cases, we decrease the width of the interval by 2.*
  - *We stop when this width becomes smaller than or equal to ε, i.e., when*
  $$\overline{x} - \underline{x} \le \varepsilon.$$
  - *Then, we take $\overline{x}$ as the desired output x_r.*

Proof. It is easy to prove that the algorithm $\mathscr{A}$ indeed solves the search under constraints problem. Indeed, increasing the previously tested value $x \le x_0$ is legitimate: since then $x + \delta \le x_0 + \delta$. By this increase, for each x_0, we will eventually find the value i for which $x_0 \le (i+1) \cdot \delta$ – namely, $i = \left\lceil \frac{x_0}{\delta} \right\rceil - 1$. Then, by induction, we can prove that on each step of the bisection process, we indeed have $\underline{x} < x_0 \le \overline{x}$. And if $\underline{x} < x_0 \le \overline{x}$ and $\overline{x} - \underline{x} \le \varepsilon$, then indeed $x_0 \le x_r = \overline{x} \le \underline{x} + \delta < x_0 + \varepsilon$.

Optimality if also easy to prove: indeed, the algorithm $\mathscr{A}$ takes $\frac{x_0}{\delta} + \text{const}$ steps, where the constant – approximately equal to $\log_2\left(\frac{\delta}{\varepsilon}\right)$ – covers the bisection part. Let us show that other algorithms A cannot use fewer steps.

Indeed, if v is the largest value for which we have already checked that $v < x_0$, then, at the next test, we cannot use the value $x > v + \delta$. Indeed, in this case, we have $v < x - \delta$ so for any x_0 from the interval $(v, x - \delta)$, we have $v < x_0 < x - \delta$ and thus, $x > x_0 + \delta$. So, for this x_0, the checking algorithm c is not applicable.

Thus, at each step, we cannot increase the tested value x by more than δ in comparison with the previously tested value. So, to get to a value $x \ge x_0$ – which is our goal – we need to make at least $\frac{x_0}{\delta}$ calls to the checking algorithm c.

The proposition is proven.

Comment. This is exactly what the both doctors did, the difference is that:

- the first doctor used $\delta = 200$, while
- the second doctor used a much smaller value $\delta = 25$.

Acknowledgments

This work was supported in part by the US National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science) and HRD-1242122 (Cyber-ShARE Center of Excellence).

References

1. Th. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, MIT Press, Cambridge, Massachusetts, 2009.