二分探索

1 導入

二分探索で最重要なのは、**真ん中を見ることそのものではなく、1 回の比較で半分を捨てられる条件を持っていること**です。

その条件が、配列が整列していることです。順序があるからこそ、中央の値と比べた結果から、左半分か右半分のどちらかをまとめて除外できます。この講義では、その論理を不変条件として整理します。

2 用語と定義

まず主役になる用語をそろえます。

二分探索 (Binary search) とは、整列済みの配列で、探索範囲を半分ずつ狭めながら目標値を探す方法です。

整列済み (sorted) とは、$A_0 \le A_1 \le \dots \le A_{n-1}$ が成り立つことです。

探索区間 (search interval) は、まだ x があるかもしれない範囲です。この講義では $[l, r)$ を使い、候補は $A_l, A_{l+1}, \dots, A_{r-1}$ とします。

中央 (middle) の添字は $m = \left\lfloor \frac{l+r}{2} \right\rfloor$ とします。

不変条件 (invariant) とは、反復の各段階でずっと保たれる条件です。

3 方針

二分探索では、まず「もし x が存在するなら、必ず探索区間 $[l, r)$ の中にある」という不変条件を立てます。

つぎに中央 A_m と x を比べます。配列が整列済みなので、

- $x < A_m$ なら m 以降は全部大きすぎる
- $x > A_m$ なら m 以前は全部小さすぎる

と言えます。つまり、中央を見る理由は「半分を捨てる証拠を得るため」です。

4 導出

4.1 1. 探索区間をどう持つか

最初は、もし x が存在するなら配列全体のどこかにあります。したがって初期状態は $l = 0, r = n$ として、候補を $[0, n)$ にします。

このときの不変条件は、

$$\boxed{\text{もし } x \text{ が存在するなら、その位置は常に } [l, r) \text{ の中にある}}$$

です。

4.2 2. 中央との比較で何が捨てられるか

$m = \left\lfloor \frac{l+r}{2} \right\rfloor$ として A_m を見ます。

まず $A_m = x$ なら、その場で探索成功です。

つぎに $x < A_m$ の場合を考えます。配列は整列済みなので、$A_m, A_{m+1}, \ldots, A_{r-1}$ はどれも A_m 以上です。したがって、これらは x にはなりえません。よって候補は $[l, m)$ だけに絞れます。つまり $r = m$ としてよいです。

反対に $x > A_m$ なら、$A_l, A_{l+1}, \ldots, A_m$ はどれも A_m 以下です。したがって、これらは x にはなりえません。よって候補は $[m+1, r)$ に絞れます。つまり $l = m + 1$ としてよいです。

どちらの場合も、不変条件は保たれたまま、探索区間の長さ $r - l$ が縮みます。

4.3 3. なぜ必ず終わるのか

$A_m = x$ でない限り、探索区間は $[l, m)$ または $[m+1, r)$ に変わります。どちらも元の $[l, r)$ より真に短い区間です。

したがって $r - l$ は反復のたびに減ります。よって有限回でかならず $l = r$ になり、探索区間は空になります。このときは x が存在しないと結論できます。

4.4 4. 計算量

探索区間の長さは、おおむね 1 回ごとに半分になります。したがって、n 個の候補を 1 個まで絞るのに必要な回数はおよそ $\log_2 n$ 回です。

よって二分探索の計算量は $O(\log n)$ です。

5 どこまで成り立つか

二分探索が使えるのは、配列や探索対象に順序があり、その順序を使って「こちら側は全部不可能」と言えるときです。

したがって、整列していない配列にはそのままでは使えません。また、比較の結果から半分を捨てられない問題にも使えません。

つまり、前提条件は線形探索より強いですが、その代わりに計算量は $O(\log n)$ まで下がります。

6 見分け方

- 比較の結果だけで、ある範囲をまとめて不可能にできるなら、二分探索を疑います。
- 配列が整列している、あるいは「単調に真偽が切り替わる」構造があるなら、半分を捨てられる可能性があります。

- 順序はあるが、比較してもどちら側を全部捨ててよいか言えないなら、二分探索ではなく別の方法を考えます。

7 最終形

整列済み配列に対して、候補を半開区間 $[l, r)$ で持ちます。

- 初期値は $[l, r) = [0, n)$ です。
- 中央 $m = \left\lfloor \frac{l+r}{2} \right\rfloor$ を取ります。
- $A_m = x$ なら成功です。
- $x < A_m$ なら $r = m$ にします。
- $x > A_m$ なら $l = m + 1$ にします。
- $l = r$ になったら「存在しない」と結論します。

したがって二分探索は

$$\boxed{\text{整列によって不可能な半分を捨て続ける探索}}$$

であり、計算量は $\boxed{O(\log n)}$ です。

8 一言でいうと

- 中央を見るのは、中央が特別だからではなく、半分を捨てる根拠を得やすいからです。
- 本体は、「もし x があるなら $[l, r)$ の中にある」という不変条件です。
- 整列という前提条件を使って候補集合を半分ずつ削る、それが二分探索です。

9 関連リンク

→ 講義 **線形探索** lecture information algorithm
https://study.bem130.com/lecture/information/algorithm/search/線形探索-講義/