Quicksort

Randomized Quicksort

1. Pick a random element as pivot
2. Split array into sub-arrays LESS, EQUAL, GREATER
3. Recursively sort sub-arrays

Analysis of Randomized QS

Assume all elements are distinct.

Thm The expected # of comparisons made by RQS is $\underbrace{2n \ln n}$.

Pf When we pick a pivot, perform $(n-1)$ comparisons to split array.

Depending on pivot choice, we have following possibilities.

$|LESS| = 0, \quad |GREATER| = n-1$

OR $\quad$ " $= 1, \quad$ " $= n-2$

" $= 2, \quad$ " $= n-3$

$\vdots \qquad\qquad \vdots$

$|LESS| = n-1, \quad$ " $= 0$

$$T(n) = \underbrace{n-1}_{\text{splitting}} + \underbrace{\frac{1}{n}}_{\text{prob. weightage}} \left(\sum_{i=0}^{n-1} \left(\underbrace{T(i)}_{\text{Rec. on } |L|} + \underbrace{T(n-i-1)}_{\text{Rec. on } |G|} \right) \right)$$

$$E(X) = \sum_{x \in X} p(x) f(x)$$

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

Guess the solution is $cn \ln n$, for some constant c.

$$\boxed{\sum_{i=1}^{n-1} f(i) \le \int_1^n f(x)\,dx,}$$ $f(x)$ is an increasing function.

$$\underbrace{\int cx \ln x \, dx}_{} = \frac{cx^2}{2} \ln x - \frac{cx^2}{4}$$

We guess $T(i) \le c i \ln i$ for $i \le n-1$.

Base Case: $T(1) = 0$

$$T(n) \le n-1 + \frac{2}{n} \sum_{i=1}^{n-1} c i \ln i$$

$$\le (n-1) + \frac{2}{n} \int_1^n cx \ln x \, dx$$

$$= (n-1) + \frac{2}{n}\left(\frac{c}{2} n^2 \ln n - \frac{cn^2}{4} + \frac{c}{4} \right)$$

$$\le c n \ln n \quad (c = 2)$$

$$\boxed{= 2n \ln}$$

Selection: Given an unsorted array, how quickly can we find the median. — ($\frac{n}{2}^{th}$ element of sorted list)

→ general problem ⟶ find the k^{th} element in the sorted list.

Randomized Version

Sort & output k^{th} element its $O(n \lg n)$

Want linear solution.

Observation

In RQS after partitioning we can tell whether the k^{th} element is in LESS OR GREATER subarray, based on their sizes.

e.g. want 87^{th} elelement, $|L| = 200$, then

L must contain 87^{th} element, can discard G.

Randomized Select

Input: array A of size n & int $k \le n$

1. Pick pivot p randomly from A.

2. Split A into L & G, by comparing with p

3. if $|L| = k-1$ output p
   if $|L| > k-1$ output Random Select (L, k)
   if $|L| < k-1$, " " " $(G, k-|L|-1)$

$$T(n) \le n-1 + \frac{1}{n/2} \sum_{i=n/2}^{n-1} T(i)$$

<u>Intuition:</u> Split interval $[0, 1]$ uniformly random into two, to get random variable u.

<u>Define Y</u> $= \max(u, 1-u)$

$$E(Y) = \int_0^1 E(Y \mid u)\, du = \int_0^{.5} (1-u)\, du + \int_{.5}^1 u \cdot du$$

$$= \frac{3}{4}$$