COS 488
Eric Wu - ezwu@
Feb 8, 2017

5/5

PS1 - Q2

Let's denote by S_N the expected number of subarrays of size 2 or less encountered when sorting a random file of size N. If quicksort is called on an array of size $N = 0 \text{or} N = 1$, then it terminates immediately, having encountered 1 subarray of size 2 or less. This gives the base cases $S_0 = 1, S_1 = 1$. If quicksort is called on an array of size 2, then it will have encountered 1 subarray of size 2, and then recursively call quicksort on one subarray of size 0 and one subarray of size 1, giving $S_2 = 3$. For any $N > 2$, quicksort no longer encounters a subarray of size 2 or less on the initial call, and can be defined recursively as:

$$S_N = \frac{1}{N} \sum_{1 \le j \le N} (S_{j-1} + S_{N-j}) \qquad \text{for } N > 2 \tag{1}$$

Using the symmetry of the sum, we get:

$$S_N = \frac{2}{N} \sum_{1 \le j \le N} S_{j-1} \qquad \text{for } N > 2 \tag{2}$$

Multiplying by N and subtracting the resulting equation for $N - 1$

$$NS_N - (N-1)S_{N-1} = 2S_{N-1} \qquad \text{for } N > 3 \tag{3}$$

Rearranging and dividing by $N(N+1)$ gives:

$$\frac{S_N}{N+1} = \frac{S_{N-1}}{N} \qquad \text{for } N > 3 \tag{4}$$

and iterating gives:

$$\frac{S_N}{N+1} = \frac{S_3}{4} \qquad \text{for } N > 3 \tag{5}$$

Using $S_3 = \frac{2}{3} \sum_{1 \le j \le 3} S_{j-1} = \frac{10}{3}$, we have that:

$$\frac{S_N}{N+1} = \frac{5}{6} \qquad \text{for } N > 3 \tag{6}$$

$$S_N = \frac{5}{6}(N+1) \qquad \text{for } N > 3 \tag{7}$$

For $N = 3$ we have that $\frac{5}{6}(3+1) = \frac{10}{3}$ and so the general relation is:

$$S_N = \frac{5}{6}(N+1) \qquad \text{for } N > 2 \tag{8}$$

with the base cases $N = 0, 1, 2$ defined as above.