Quicksort Aleatorizado

André Vignatti

Quicksort:

- Pior caso: $\Theta(n^2)$
- Na prática: $\Theta(n \log n)$, ganha do Mergesort (que não tem pior/melhor caso)

Análise do caso médio e da aleatorizado são quase iguais.

- **Caso médio** assume **todas entradas equiprováveis**: mentira!

Usaremos **aleatorização** no QuickSort.

Ideia: Escolher o pivô **aleatoriamente**.

ParticioneAle(v, a, b)
$i \leftarrow$ **random**(a, b)
Troca(v, i, b)
devolva Particione(v, a, b)

É **esperado**, dividir o vetor de maneira **bem balanceada**.

QuickSortAle(v, a, b)
`se` $a < b$
  `entao` $m \leftarrow$ ParticioneAle(v, a, b)
  QuickSortAle$(v, a, m - 1)$
  QuickSortAle$(v, m + 1, b)$

Rrecorrência: o custo do algoritmo é no Particiona

- A **operação fundamental** do Particiona (veja o pseudo=código) é a comparação (if).

Então, vamos contar **todas** as comparações feitas pelo algoritmo.

Supomos elementos **distintos**

Seja $z_1 < z_2 < \ldots < z_n$ os elementos ordenados

Seja X_{ij} v.a. tal que: z_j:

$$X_{ij} = \begin{cases} 1 & \text{se } z_i \text{ é comparado com } z_j \\ 0 & \text{caso contrário.} \end{cases}$$

Fatos Importantes:

1. Comparações só ocorrem com o pivô.
2. Após `Particione`, o pivô não participa de outras comparações até o fim da execução.
3. Um par de elementos é comparado no **máximo** uma vez!
4. Assim, $X_{ij} = X_{ji}$. Para evitar contar duas vezes, vamos somente contar X_{ij} se $i < j$.

Então, o **número total de comparações** X **é**

$$X = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}.$$

e o número **esperado** de comparações é $E[X]$. Pela **linearidade da esperança**:

$$E[X] = E\left[\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}\right] = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E[X_{ij}]$$

Como X_{ij} é uma variável aleatória binária,

$$\begin{aligned} E[X_{ij}] &= 0 \cdot \Pr(z_i \text{ não ser comparado com } z_j) + \\ & \quad 1 \cdot \Pr(z_i \text{ ser comparado com } z_j) \\ &= \Pr(z_i \text{ ser comparado com } z_j) \end{aligned}$$

Então,

$$E[X] = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \Pr(z_i \text{ ser comparado com } z_j)$$

Agora basta encontrar a probabilidade de dois elementos serem comparados.

Considere a escolha do pivô e a comparação entre z_i e z_j:

$$\underbrace{z_1, z_2, \ldots, z_{i-1}}_{\text{posterga}}, \mathbf{z_i}, \underbrace{z_{i+1}, \ldots, z_{j-1}}_{\textbf{não comp.}}, \mathbf{z_j}, \underbrace{z_{j+1}, \ldots, z_n}_{\text{posterga}}$$

- Se algum **azul** é escolhido como pivô, z_i e z_j **nunca mais** serão comparados **(pois ficarão em particões diferentes)**
- Para serem comparados, ou z_i ou z_j devem ser escolhidos como pivô **ANTES** de algum **azul**.

Seja $Z_{ij} = \{z_i, z_{i+1}, \ldots, z_j\}$. Assim,

$$
\begin{aligned}
&Pr(z_i \text{ ser comparado com } z_j) \\
&\qquad = \Pr(z_i \text{ ou } z_j \text{ é escolhido como pivô primeiro em } Z_{ij}) \\
&\qquad = \Pr(z_i \text{ é escolhido como pivô primeiro em } Z_{ij}) + \\
&\qquad\quad\ \Pr(z_j \text{ é escolhido como pivô primeiro em } Z_{ij}) \\
&\qquad = \frac{1}{j-i+1} + \frac{1}{j-i+1} \\
&\qquad = \frac{2}{j-i+1}
\end{aligned}
$$

(A segunda linha é verdade pois os dois eventos são mutuamente exclusivos)

Substituindo na equação para $E[X]$,

$$
\begin{aligned}
\mathbf{E[X]} &= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E[X_{ij}] \\
&= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{2}{j-i+1} \\
&= \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} \frac{2}{k+1} \\
&< \sum_{i=1}^{n-1} \sum_{k=1}^{n} \frac{2}{k} \\
&= \text{(num. harmônico, eq. A7 do CLRS)} \sum_{i=1}^{n-1} O(\lg n) \\
&= O(n \lg n)
\end{aligned}
$$

O consumo de tempo esperado pelo QuickSortAle é $O(n \lg n)$.

O limitante de melhor caso visto anteriormente (dividir "exatamente" no meio) era $T(n) = \Omega(n \lg n)$.

Conclusão: O consumo de tempo esperado pelo QuickSortAle é $\Theta(n \lg n)$.