Mar 31

Collaborative Filtering (Netflix)

Each user ≡ a ranking on movies/shows on Netflix

Hypothesis: user 1 is 'close to' user 2 if user 1's ranking is 'close to' user 2's ranking.

Assumption: Each user ranks all movies/shows

Input: A ranking $a_1, \ldots, a_n$ (permutation on $\{1, \ldots n\}$)

Output: Number of inversions

ex. 1 3 2

Recall: (i, j) is an inversion
(1) $i < j$ AND (2) $a_i > a_j$

$a_i = 3 > a_j = 2$ ⟩ inversion

Ex. 1 $a = (1, \ldots, n)$; #inversions = 0

all pairs are NOT inversions

$a_1, \ldots, a_n$ being sorted ≡ #inversions = 0

(if $i < j$; $a_i = i$, $a_j = j$)

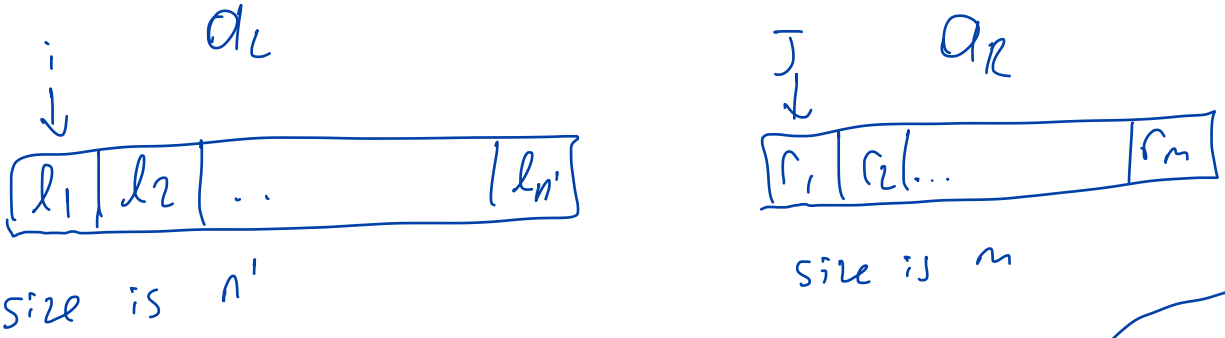
Ex. 2 $a = (n, \ldots, 1)$; $\#\text{inversions} = \binom{n}{2} = \frac{n \cdot (n-1)}{2}$

i j

$a_i > a_j$

all pairs are inversions

General 'bad case'



Goal:
Count #pairs (i,j)
$1 \le i \le n'$
$1 \le j \le m$
s.t. $l_i > r_j$

Assume a_L is sorted AND
a_R is sorted

Merge-Count (a_L, a_R)

0. $c = 0$
1. $i = 1$, $j = 1$
2. while $i \le n'$ AND $j \le m$
   If $l_i < r_j$
   $i++$
   Else
   $c += n' - i + 1$
   $j++$
3. return c