Oct 25

Collaborative filtering

Each user ≡ a ranking on movies/shows on Netflix

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

Assumption: Each user ranks all movies/shows.

Mike	Ron
1. The Office	1. Parks & Rec
2. Parks & Rec	2. The Office
3. Friends	3. Portlandia

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

Output: Number of inversions

ex. 1 3 2

↓

(3, 2) is an inversion

Recall: (i, j) is an inversion

(1) $i < j$ AND (2) $a_i > a_j$

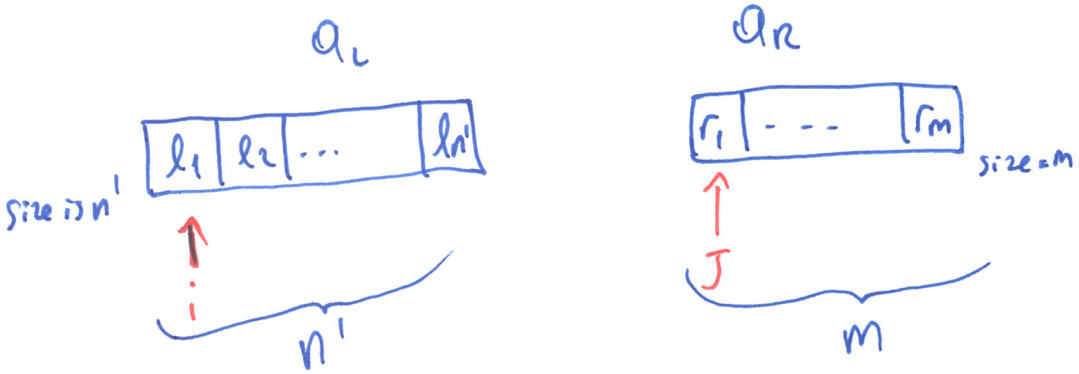
Ex 1. $a = (1, \ldots, n)$; # inversions = 0

all pairs are NOT inversions

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

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

General 'bad case'



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

Merge Count (a_L, a_R) (a_L is sorted) (a_R is sorted)

0. $c = 0$

1. $i = 1$, $j = 1$

2. while $i \leq n'$ AND $j \leq m$

    If $l_i < r_j$

        $i++$

    else

        $c += (n' - i + 1)$

        $j++$

3. return c