Oct 30

Collaborative Filtering (eg. Netflix)

Each user $\equiv$ a ranking of movies/shows on Netflix

Hypothesis: User A is "close" to user B if A's ranking is "close" to B's.

Assumption: Each user ranks ALL movies/shows on Netflix

$n=3$

User 1	User 2	User 3
① Pokemon	3	1
② Great British Baking Show	2	3
③ Beauty in Black	1	2

Counting inverses / Computing Kendall-Tau distance

input: A ranking $a_1, \ldots, a_n$ (permutation of $1 \ldots n$)

Implicit assumption: "True" / "other" ranking $1 \ldots n$

Output: The number of inversions (Kendall-Tau dist.)

Def: (i,j) is an inversion $i,j \in [n]$

if

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

Ex 1: User 2 $(a_1, a_2, a_3) = (3, 2, 1)$

$(1,2)$, $(1,3)$, $(2,3)$ are all inversions $\Rightarrow$ # inv $= 3$

Ex 2: User 3 $(a_1, a_2, a_3) = (1, 3, 2)$

only invs $(2,3)$ # inv $= 1$

Ex3 $(a_1, \ldots, a_n) = (1, 2, \ldots, n)$ # inv = ? 0

Ex4: $(a_1, \ldots, a_n) = (n, n-1, \ldots, 1)$ # inv = ?

every pair (i,j) st $i<j$ is an inversion

inv = # pairs $= \binom{n}{2} = \frac{n(n-1)}{2}$

Ex 5: $(a_1, \ldots, a_n)$ is sorted in increasing order $\Longleftrightarrow$ # inv $= 0$

Home

Ex 6: $0 \le$ # inv $\le \binom{n}{2}$