

# All Pairs Similarity via matrix multiplication-

## Assignment Q5 hint

- If  $A$  is an  $n \times m$  matrix containing *the*  $n$ -dimensional vectors (purchase histories) for  $m$  customers, we can compute all pairs similarity very straightforwardly.
- First compute all of the dot products using  $A^T A$  e.g.

$$\begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 2 & 0 \\ 3 & 3 \end{pmatrix} = \begin{pmatrix} 14 & 8 \\ 8 & 10 \end{pmatrix}$$

- Then take every element on the leading diagonal and divide all of the elements in its containing row and column by it's square root:

$$\begin{pmatrix} 14/\sqrt{14 \times 14} & 8/\sqrt{14 \times 10} \\ 8/\sqrt{10 \times 14} & 10/\sqrt{10 \times 10} \end{pmatrix} = \begin{pmatrix} 1 & 4/\sqrt{35} \\ 4/\sqrt{35} & 1 \end{pmatrix}$$

[ Divide  $i, j$  component by  
(length of vector  $i$  times  
length of vector  $j$ ) ]

- Matrix multiplication can be done in  $O(m^2n)$ , but can use Strassen's algorithm for large data.
- Dividing every element is  $O(m^2)$  so...
- All pairs similarity can be done in less than  $O(m^2n)$