

Deep Learning
Prof. Mitesh M Khapra
Department of Computer Science Engineering
Indian Institute of Technology, Madras

Mod-03-Part 2
Lecture – 10
SVD for learning word representations

(Refer Slide Time: 00:12)

$\hat{X} =$

	human	machine	system	for	...	user
human	2.01	2.01	0.23	2.14	...	0.43
machine	2.01	2.01	0.23	2.14	...	0.43
system	0.23	0.23	1.17	0.96	...	1.29
for	2.14	2.14	0.96	1.87	...	-0.13
...
user	0.43	0.43	1.29	-0.13	...	1.71

$\hat{X}\hat{X}^T =$

	human	machine	system	for	...	user
human	25.4	25.4	7.6	21.9	...	6.84
machine	25.4	25.4	7.6	21.9	...	6.84
system	7.6	7.6	24.8	18.03	...	20.6
for	21.9	21.9	0.96	21.6	...	15.32
...
user	6.84	6.84	20.6	15.32	...	17.11

similarity = 0.33

- Notice that the dot product between the rows of the matrix $W_{word} = U\Sigma$ is the same as the dot product between the rows of \hat{X}

$$\hat{X}\hat{X}^T = (U\Sigma V^T)(U\Sigma V^T)^T$$

$$= (U\Sigma V^T)(V\Sigma U^T)$$

$$= U\Sigma\Sigma^T U^T \quad (\because V^T V = I)$$

$$= U\Sigma(U\Sigma)^T = W_{word}W_{word}^T$$
- Conventionally,

$W_{word} = U\Sigma \in \mathbb{R}^{m \times k}$

 is taken as the representation of the m words in the vocabulary and

$$W_{context} = V \in \mathbb{R}^{n \times k}$$
 is taken as the representation of the context words

So, we will start from where we left yesterday ah. So, we did this whole story on starting from co occurrence matrices, we learnt how to get better word representations and the key thing there was we used SVD as a dimensionality comp reduction tool. And we came up with this neat result that you could use W_{word} as the representation of the words, it has m rows and k columns where k is very very less than the size of the vocabulary.

So, you have achieved lot of compression and you are still able to learn very meaningful representations, which you could use for several downstream tasks. What to use these for and how to use these for you will see that later maybe four lectures from now, I mean I say 4 lectures I mean 4 2 hour lectures right. So, it might be more in terms of actual lectures ah. So, we will get to that, but for now we have a way of learning representations for words.