Dimensionality Reduction / SVD

issues with X^TX when performing PCA

① memory

$p \times p$ square matrix

MNIST $n = 60{,}000$ (training examples)

$p = 28 \times 28 = 784$

$p \times p = 614{,}656$ pixel

$\times$ 8 bytes / pixel

you need 4Mb of memory to store X^TX

however, if $p = 200 \times 200$ (resolution of a LinkedIn profile picture), you would need 12Gb of memory to store X^TX

$p = 400 \times 400 \Leftrightarrow 200$ Gb

② numerical errors happen (rounding) when computing X^TX

A factorization technique called SVD allows us to perform PCA without computing X^TX

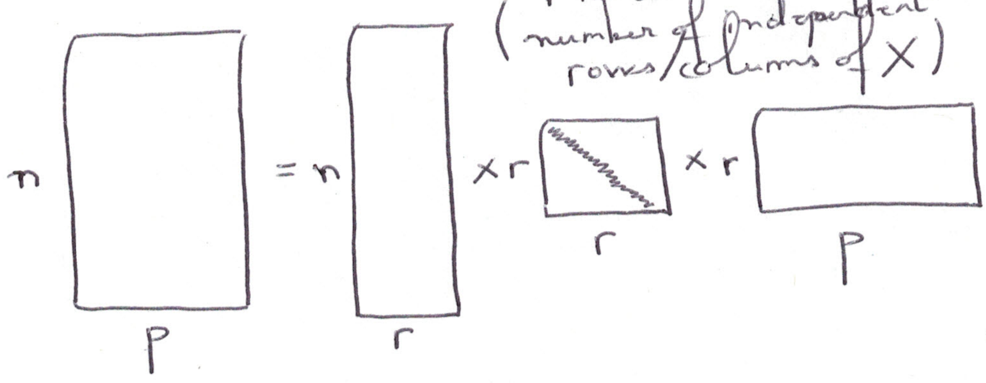
$$\underset{n \times p}{X} = \underset{n \times r}{U} \; \underset{r \times r}{\Sigma} \; \underset{p \times r}{V}^T$$

r is the rank of X (number of independent rows/columns of X)

with U and V othonormal

$U^T U = I$

$V^T V = I$



how does SVD relates to PCA?

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

$$X^T \cdot X = (V\Sigma U^T)(U\Sigma V^T)$$
$$= V\Sigma \underbrace{(U^T U)}_{I} \Sigma V^T$$
$$= V\Sigma^2 V^T$$

let's right multiply by V on both sides

$$(X^T X)V = V\Sigma^2$$

the column vectors of V are the eigenvectors of X^TX with eigenvalues λ_j^2

similarly, $XX^T = (U\Sigma \underbrace{V^T)(V}_{I}\Sigma U^T) = U\Sigma^2 U^T$

$$(XX^T)U = U\Sigma^2$$

the column vectors of U are the eigenvectors of XX^T

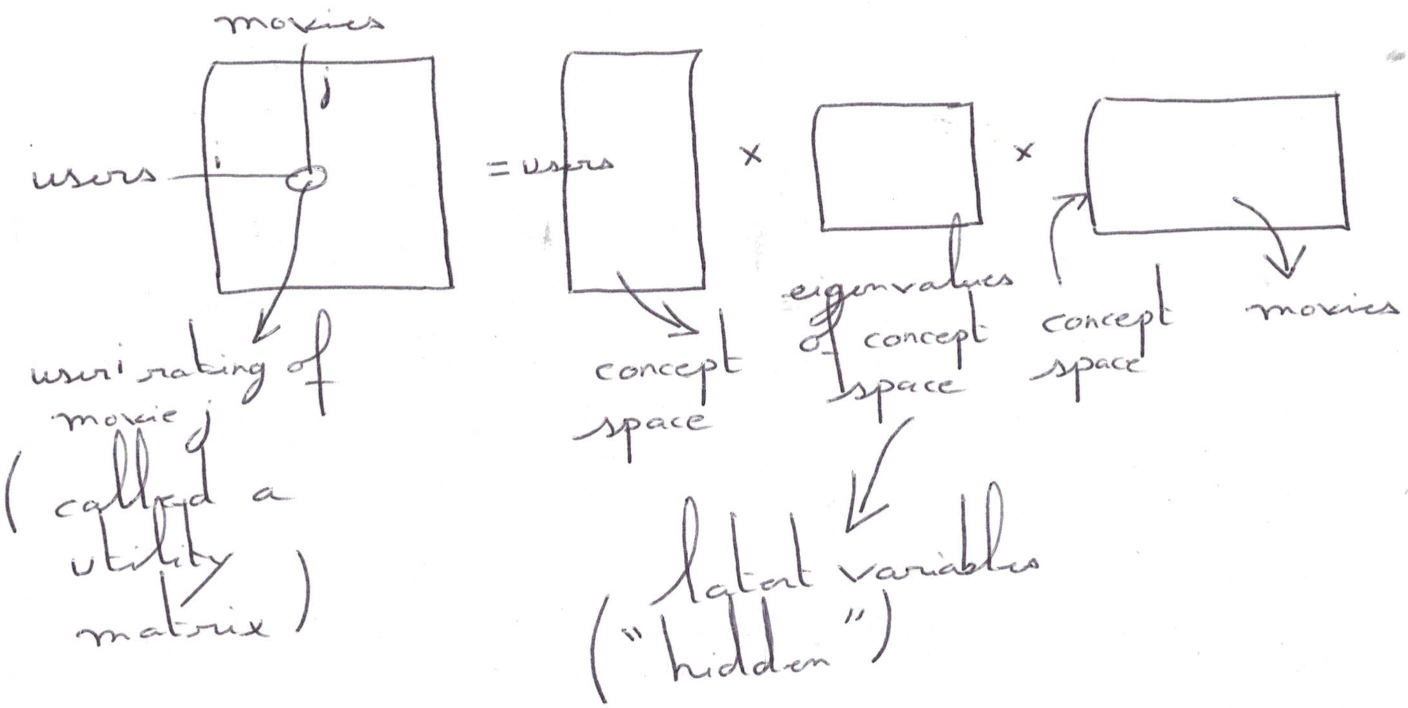
$$X = U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_r \end{pmatrix} V^T$$

columns are eigenvectors of XX^T

column vectors are eigenvectors of $X^T \cdot X$

λ_j^2 are eigenvalues for column vectors of U/V (XX^T/X^TX)

Application to latent features



	Matrix	Alien	Serenity	Casablanca	Amelie
Alice	1	2	2		
Bob	3	5	5		
Cindy	4	4	4		
Dan	5	5	5		
Emily		2		4	4
Frank				5	5
Greg		1		2	2

=

same column

Bob, Cindy, and Dan like science fiction movies

rank is 4

topic is science fiction

science fiction movies

$$
\begin{array}{c}
A\\B\\C\\D\\E\\F\\G
\end{array}
\begin{pmatrix}
-0.2 & 0 & 0.3 & -0.3\\
-0.5 & 0.1 & 0.5 & -0.5\\
-0.5 & 0.1 & -0.3 & 0.2\\
-0.6 & 0.1 & -0.4 & 0.2\\
-0.1 & -0.6 & 0.4 & 0.5\\
0 & -0.7 & -0.4 & -0.5\\
-0.1 & -0.3 & 0.2 & 0.3
\end{pmatrix}
\times
\begin{pmatrix}
13.8 & & & 0\\
 & 9.5 & & \\
 & & 1.7 & \\
0 & & & 1
\end{pmatrix}
\times
\begin{array}{c}
\begin{pmatrix}
-0.5 & -0.6 & -0.6 & -0.1 & -0.1\\
0.1 & 0 & 0.1 & -0.7 & -0.7\\
-0.8 & 0.6 & 0 & -0.1 & -0.1\\
0.4 & 0.5 & -0.8 & -0.1 & -0.1
\end{pmatrix}\\
\begin{matrix} M & \quad A & \quad S & \quad C & \quad A \end{matrix}
\end{array}
$$

romance