Singular Value Decomposition:

2/19 & 2/21

Defn: $\boxed{\underline{M}_{[m \times n]} = \underline{U}_{[m \times r]} \, \Sigma_{[r \times r]} \left(V_{[n \times r]}\right)^T}$

$\underline{M}$: Input data matrix – [m x n : e.g. m documents & n terms]

$\underline{U}$: Left singular matrix – [m x r : e.g. m documents & r concepts]

Σ : Singular values – [r x r diagonal matrix; strength of each concept]

$\underline{V}$: Right singular matrix – [n x r matrix]

What is r? how do we calculate it?

Example to Demonstrate the Interpretation of Singular Value Decomposition

	Dead Space 3	Assassin's Creed 3	Call of Duty	Fifa Soccer 13	Soccer Manager 13
User 1	1	1	1	0	0
User 2	2	2	2	0	0
User 3	1	1	1	0	0
User 4	5	5	5	0	0
User 5	0	0	0	2	2
User 6	0	0	0	3	3
User 7	0	0	0	1	1

From Matlab: r = rank(M); ⟶ r = 2

[U, Sigma, V] = svd(M);

But U is m×m (7×7), V is n×n (5×5) & Sigma is m×n (7×5)

∴ we should keep the columns & rows that correspond to r

i.e. $\underline{U} = \underline{U}(:, 1{:}r)$;

$\underline{V} = \underline{V}(:, 1{:}r)$;

Σ = Sigma [Sigma(1,1) Sigma(1,2); Sigma(2,1) Sigma(2,2)]

Solution:

$$\underline{U} = \begin{bmatrix} 0.18 & 0 \\ 0.36 & 0 \\ 0.18 & 0 \\ 0.90 & 0 \\ 0 & 0.53 \\ 0 & 0.80 \\ 0 & 0.27 \end{bmatrix} \longrightarrow$$

connects people to concepts; e.g. User 1: - Likes Action Games
- but rates them Less than Users 2 - 4
- Does not Like (has not rated) any Soccer games

$$\Sigma = \begin{bmatrix} \overset{\sigma_1}{9.64} & 0 \\ 0 & \underset{\sigma_2}{5.29} \end{bmatrix} \longrightarrow \text{strength of each concept}$$

$$\underline{V}^T = \begin{bmatrix} 0.58 & 0.58 & 0.58 & 0 & 0 \\ 0 & 0 & 0 & 0.71 & 0.71 \end{bmatrix} \longrightarrow \text{connects games to concepts}$$

Note: Both $\underline{U}$ and $\underline{V}$ are column orthonormal
- each of their columns is a unit vector
- dot product of any 2 columns is zero
- $U^T U$ = identity Matrix
- $V^T V$ = identity Matrix

Querying Using Concepts: Read book P. 390 – 391; great explanation (very simple to read)

→ we did talk about briefly in class [relevant to ch. 9 in book]

Computing SVD of a Matrix: Last class, we ended here:

<u>Defn</u>: $\underline{M} = \underline{U}\ \underline{\Sigma}\ \underline{V}^T$

<u>Approach</u>: Calculate $M^T M$ and $M M^T$ $\xrightarrow{\text{bec}}$ helpful in understanding how to obtain the R.H.S of the above equation

$$\underline{M}^T = (\underline{U}\,\underline{\Sigma}\,\underline{V}^T)^T = \underline{V}\,\underline{\Sigma}\,\underline{U}^T$$

$$\underline{M}^T\underline{M} = (\underline{V}\,\underline{\Sigma}\,\underline{U}^T)\,\underline{U}\,\underline{\Sigma}\,\underline{V}^T = \underline{V}\,\underline{\Sigma}^2\,\underline{V}^T$$

$$\underline{M}^T M\,\underline{V} = \underline{V}\,\underline{\Sigma}^2\,\underline{V}^T\,\underline{V}$$

$$\boxed{\underline{M}^T\underline{M}\,\underline{V} = \underline{V}\,\underline{\Sigma}^2} \quad \Rightarrow \underline{A}\,\underline{V} = \underline{V}\,\underline{\Sigma}^2$$

$$M M^T =$$

Dimensionality Reduction: – Suppose you want to represent M (very large) by its SVD components, these matrices may also be very large to store conveniently.

What shall we do? :)

→ how many concepts to keep? → Energy

→ $\text{Energy} = \dfrac{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_j^2}{\sum_{i=1} \sigma_i^2}$

→ zero out the rows and Columns corresponding to the concepts that we are taking out.

→ $U = \begin{bmatrix} 0.18 \\ 0.36 \\ 0.18 \\ 0.9 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $\Sigma = [9.64]$, $\underline{V}^T = \begin{bmatrix} 0.58 & 0.58 & 0.58 & 0 & 0 \end{bmatrix}$

Finding the Rank of the Matrix : → # ind rows or columns in that matrix

→ $n \times n$ matrix
↳ Rank : $[0, n]$

→ $m \times n$ matrix ; assume $n < m$
Rank : $[0, n]$

Properties for indep. rows:

1] $R_i \neq R_j \times \underline{\underline{a}}$ ⟶ anyconstant (for columns too)

2] The row/column has to have @ least 1 non-zero value

3] A row/column should not be a linear combination of another row/column

Ex: $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \longrightarrow r = ?$

row echleon form of the matrix:

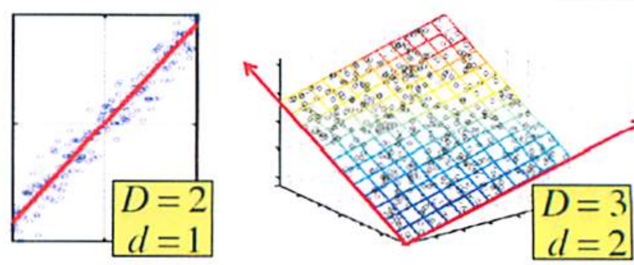
1) The first element of the matrix 1.
2) essentially to have an upper triangular matrix

$$\begin{array}{r} \\ -4R_1 + R_2 \rightarrow \\ -7R_1 + R_3 \end{array} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{bmatrix} \xRightarrow{-2R_{2new} + R_{3new}} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{bmatrix}$$

$r = 2$

Conclusions: Dimension Reduction – What it is and Why? 2

- We talked about two techniques (PCA and SVD)
  - There exists many others
- High dimensional data implies that we have many features → So why do we want to reduce it?
  - Discover hidden correlations/topics
  - Remove redundant and noisy features
  - Easier storage & processing
  - Interpretation/ visualization



- **Assumption:** Data lies on or near a low d-dimensional subspace
- Axes of this subspace are effective representation of the data

$$\Sigma = \begin{bmatrix} 9 & 0 \\ 0 & 5 \end{bmatrix}$$

$$\Sigma \times \Sigma = \underset{(2\times 2)}{\begin{bmatrix} 9 & 0 \\ 0 & 5 \end{bmatrix}} \underset{(2\times 2)}{\begin{bmatrix} 9 & 0 \\ 0 & 5 \end{bmatrix}}$$

$$= \begin{bmatrix} 9^2 & 0 \\ 0 & 5^2 \end{bmatrix} =$$

$$\frac{9.64^2}{9.64^2 + 5.29^2}$$ 1 concept

$$\frac{9.64^2 + 5.29^2}{9.64^2 + 5.29^2}$$ 2 concepts