CIS 635 Knowledge Discovery & Data Mining

- Singular Value Decomposition (SVD)

Singular Value Decomposition (SVD)

- Another Linear Dimensionality Reduction Technique
- Different than PCA

PCA steps – Recall

1. Let assume D_{mxn} is a given data matrix
2. Apply standard scalar (normalization)
3. Estimate Covariance (matrix), A_{nxn}
4. Compute Eigenvalues and Eigenvectors of the Covariance Matrix by solving the linear dynamical system at the right

If ***A*** is a *nxn* matrix, solving this linear dynamical system will give *n* **eigenvalues**, and n associated *n* **eigenvectors**

$$\boxed{AX = \lambda X}$$

$$AX - \lambda X = 0$$

or

$$(A - \lambda I)\,X = 0$$

SVD

- *A* is your Data Matrix (Transposed)
- **SVD** decomposes *A* into 3 matrices: ***U, S, V***

$$A_{m\times n} = U_{m\times m} S_{m\times n} V^T_{n\times n}$$

$$\underset{m \times n}{\overset{\text{A}}{\begin{pmatrix} x_{11} & x_{12} & & x_{1n} \\ & & \ddots & \\ x_{m1} & & & x_{mn} \end{pmatrix}}} = \underset{m \times m}{\overset{\text{U}}{\begin{pmatrix} u_{11} & & u_{m1} \\ & \ddots & \\ u_{1m} & & u_{mm} \end{pmatrix}}} \underset{m \times n}{\overset{\text{S}}{\begin{pmatrix} \sigma_1 & & 0 \\ & \sigma_r & \\ 0 & & 0 \end{pmatrix}}} \underset{n \times n}{\overset{\text{V}^\text{T}}{\begin{pmatrix} v_{11} & & v_{1n} \\ & \ddots & \\ v_{n1} & & v_{nn} \end{pmatrix}}}$$

SVD

- *A* is your Data Matrix (Transposed)
- **SVD** decomposes *A* into 3 matrices: ***U, S, V***
- Say we are given,

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

$$\boxed{A_{m\times n} = U_{m\times m} S_{m\times n} V^T_{n\times n}}$$

$$\begin{array}{cccc} \text{A} & \text{U} & \text{S} & \text{V}^\text{T} \\ \begin{pmatrix} x_{11} & x_{12} & & x_{1n} \\ & & \ddots & \\ x_{m1} & & & x_{mn} \end{pmatrix} = & \begin{pmatrix} u_{11} & & u_{m1} \\ & \ddots & \\ u_{1m} & & u_{mm} \end{pmatrix} & \begin{pmatrix} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_r & \\ & & & \ddots \\ 0 & & & 0 \end{pmatrix} & \begin{pmatrix} v_{11} & & v_{1n} \\ & \ddots & \\ v_{n1} & & v_{nn} \end{pmatrix} \\ m \times n & m \times m & m \times n & n \times n \end{array}$$

SVD

- *A* is your Data Matrix (Transposed)
- **SVD** decomposes *A* into 3 matrices: ***U, S, V***
- Say we are given,

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

SVD solves this linear system (mxm)

$$(AA^T - I)X = 0$$

$$AA^T = \begin{pmatrix} 17 & 8 \\ 8 & 17 \end{pmatrix}$$

eigenvalues: $\lambda_1 = 25, \lambda_2 = 9$

eigenvectors

$$u_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \quad u_2 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$$

SVD

- *A* is your Data Matrix (Transposed)
- **SVD** decomposes *A* into 3 matrices: ***U, S, V***
- Say we are given,

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

And

SVD solves this linear system (nxn)

$$(A^T A - I)X' = 0$$

$$A^T A = \begin{pmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{pmatrix}$$

eigenvalues: $\lambda_1 = 25, \lambda_2 = 9, \lambda_3 = 0$

eigenvectors

$$v_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix} \quad v_2 = \begin{pmatrix} 1/\sqrt{18} \\ -1/\sqrt{18} \\ 4/\sqrt{18} \end{pmatrix} \quad v_3 = \begin{pmatrix} 2/3 \\ -2/3 \\ -1/3 \end{pmatrix}$$

SVD

- *A* is your Data Matrix (Transposed)
- **SVD** decomposes *A* into 3 matrices: ***U, S, V***
- Say we are given,

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

$$A_{m\times n} = U_{m\times m} S_{m\times n} V_{n\times n}^T$$

$$A = USV^T = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \\ 2/3 & -2/3 & -1/3 \end{pmatrix}$$

[Full derivation]()

SVD

Notebook Link