Singular Value Decomposition

이전에 $A = S\Lambda S^{-1}$ for square matrix 를 공부했다. 근데 그럼 square가 아닌 행렬의 경우는 ? 이에 관한 내용이 SVD 이다. A 를 $m*n$ 행렬이라고 하고 공부해보자. SVD를 하면 아래와 같이 3개의 행렬로 분해된다.

$$A_{m*n} = U_{m*m} \sum_{m*n} V_{n*n}^T$$

<복습 for SVD>

- when A_{m*n},
- $C(A) \in R^m \perp N(A^T) \in R^m$
- $C(A^T) \in R^n \perp N(A) \in R^n$

(1) First find the eigenvalues & eigenvectors of A^TA :

- $A^TA : n*n$, $symmetric$
- $\vec{e}^TA^TA\vec{e} = \lambda\vec{e}^T\vec{e}$ by ($A^TA\vec{e} = \lambda\vec{e}$)

$$\lambda = \frac{\|A\vec{x}\|^2}{\|\vec{x}\|^2} \geq 0$$

  A^TA의 고유값들은 0이상 이다.
- assume that there are r개 non-zero eigenvalues
  $\sqrt{\lambda_i} = \sigma_i \ (i = 1, 2, ..., r)$: singular value 라고 하고 이에 해당하는 고유벡터들은 $v_1, v_2, ...v_r \in R^n$ 서로 orthogonal(orthonormal)하다. 그러면 singular value로 이루어진 대각행렬은 아래와 같이 표현할 수 있다.

$$\sum_{r*r} = \begin{bmatrix} \sigma_1 & 0 & ... & 0 \\ 0 & \sigma_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \sigma_n \end{bmatrix}$$

$$\sum_{m*n} = \begin{bmatrix} \sum_{r*r} & 0 \\ 0 & 0 \end{bmatrix}$$

(2) V 행렬 구하기

- $V_1 = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & .... & \vec{v}_r \end{bmatrix} \in R^n$ 라고 하자.
- $\vec{v}_i$는 orthogonal하기 때문에 independent하고 이들은 $C(A^T)$ 즉, row space basis 라는 것을 알 수 있다.
- 나머지 고유값 $\lambda_{r+1} = \lambda_{r+2} = ... = \lambda_n = 0$ 에서 $A^TA\vec{x}_j = \lambda_j\vec{x}_j = 0$ 이므로 $\vec{x}_j = N(A^TA) = N(A)$ 이다. 이 고유값들에 해당하는 고유벡터들 $\vec{v}_{r+1}, \vec{v}_{r+2}, ..., \vec{v}_n$ 은 symmetric matrix의 고유벡터이므로 orthogonal(orthonormal) 하다. 이 벡터들은 $N(A)$의 basis 임을 알 수 있다. $V_2 = \begin{bmatrix} \vec{v}_{r+1} & \vec{v}_{r+2} & .... & \vec{v}_n \end{bmatrix} \in R^n$ 이고
- $V_1, \ V_2$는 서로 orthogonal하다. ($C(A^T) \perp N(A)$ 이므로) 이 내용을 토대로 V를 찾았다.

$$V_{n*n} = [V_1 \ V_2]$$

(3) U 행렬 구하기

- $AV = U\sum$의 형태로 U를 구하자. ($V^T = V^{-1}$)

$$AV = A[\vec{v}_1\ \vec{v}_2\ ...\ \vec{v}_r\ \vec{v}_{r+1} ...\ \vec{v}_n] = [A\vec{v}_1\ A\vec{v}_2\ ...\ A\vec{v}_r\ \vec{0}\ ...\ \vec{0}]$$

$$U\sum = [\vec{u}_1\ \vec{u}_2\ ...\ \vec{u}_r\ \vec{u}_{r+1} ...\ \vec{u}_n] \begin{bmatrix} \sum_{r*r} & 0 \\ 0 & 0 \end{bmatrix} = [\sigma_1\vec{u}_1\ \sigma_2\vec{u}_2\ ...\ \sigma_r\vec{u}_r\ \vec{0}\ ...\ \vec{0}]$$

위에 따라서 $A\vec{v}_i = \sigma\vec{u}_i$ 이고 따라서 $\vec{u}_i = \frac{A\vec{v}_i}{\sigma_i} \in R^m$ $(i = 1, 2, ...r)$

- $\vec{u}_i$가 orthogonal?
  $u_i^T u_j = \frac{\vec{v}_i^T A^T A\vec{v}_j}{\sigma_i\sigma_j} = \frac{\vec{v}_i^T \lambda_j \vec{v}_j}{\sigma_i\sigma_j}$ 따라서 $i = j$이면 1이고 $i \neq j$이면 0의 값을 갖기 때문에 orthogonal(orthonormal)하고 $\in R^m$이므로 이 벡터들은 $C(A)$의 basis 라고 할 수 있다.

$$U_1 = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & .... & \vec{u}_r \end{bmatrix}$$

- $\vec{u}_{r+1}, \vec{u}_{r+2}, ..., \vec{u}_m \in R^m$ 이고 이들은 left null space $N(A^T)$의 basis 이다. (위에서 구한 V와 비슷한 맥락)

$$U_2 = \begin{bmatrix} \vec{u}_{r+1} & \vec{u}_{r+2} & .... & \vec{u}_n \end{bmatrix}$$

- 최종 U

$$U = [U_1\ U_2]$$