SVD 이해를 위한 지식

참조 : https://darkpgmr.tistory.com/106

1. 선형변환
2. 고유값, 고유벡터
3. 고유값 고유벡터의 계산
4. 대칭행렬과 고유값 분해
5. 직교, 정규직교, 직교행렬
6. SVD

선형변환

벡터가 행렬을 통해 다른 벡터로 바뀌는 과정

선형변환의 성질

$$T(\vec{a}+\vec{b}) = T(\vec{a}) + T(\vec{b})$$

$$T(c\vec{d}) = cT(\vec{d})$$

$$T(\begin{bmatrix} x \\ y \end{bmatrix}) = T(x\begin{bmatrix} 1 \\ 0 \end{bmatrix}) + T(y\begin{bmatrix} 0 \\ 1 \end{bmatrix}) = xT(\begin{bmatrix} 1 \\ 0 \end{bmatrix}) + yT(\begin{bmatrix} 0 \\ 1 \end{bmatrix})$$

$$i_{new} = T(\begin{bmatrix} 1 \\ 0 \end{bmatrix}), j_{new} = T(\begin{bmatrix} 0 \\ 1 \end{bmatrix})$$

T가 선형변환이라면 벡터 [x y]는 선형 변환 후에 새로운 기저 벡터 i_new ,j_new의 x배와 y배의합으로 표현되어야 함

$$A = \begin{bmatrix} 2 & -3 \\ 1 & 1 \end{bmatrix}, \vec{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$A\vec{x} = \begin{bmatrix} 2 & -3 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

$$새로운기저벡터 A\begin{bmatrix} 1 \\ 0 \end{bmatrix} 와 A\begin{bmatrix} 0 \\ 1 \end{bmatrix} 가생성$$

$$xA\begin{bmatrix} 1 \\ 0 \end{bmatrix} + yA\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$x\begin{bmatrix} 2 & -3 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} + y\begin{bmatrix} 2 & -3 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2x-3y \\ x+y \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

$$x = y = 1$$

위의 식에서 x와 y는 각각 1이기 때문에 선형변환의 성질을 만족한다

고유값, 고유벡터

행렬A를 선형변환으로 봤을 때, 선형변환 A에의한 변환 결과가 자기 자신의 상수배가 되는 0이 아닌 벡터

$$Av = \lambda v$$

$$v = \text{고유벡터}, \lambda = \text{고유값}$$

v 가[v1,v2,v3....vn]의 벡터들로 구성되어 있을 때 고유값과 고유벡터의 곱의 열벡터는 열벡터와 고유값들을 대각 원소로 하는 대각행렬의 곱으로 나타난다

$$A\,[\,v_1 \quad v_2 \ldots vn\,] = [\,\lambda_1 v_1 \cdot \lambda_2 v_2 \ldots \lambda_n v_n\,] = [\,v_1 \quad v_2 \ldots v_n\,] \begin{bmatrix} \lambda_1 & & 0 \\ & \ldots & \\ 0 & & \lambda_n \end{bmatrix}$$

따라서 행렬 A의 고유벡터들을 열벡터로 하는 행렬을 P, 고유값들을 대각원소로 하는 대각 행렬을 L 이라고 했을 때 다음 식이 성립

$$AP = PL \quad A = PLP^{-1}$$

자신의 고유 벡터들을 열벡터로 하는 행렬과 고유값을 대각원소로 하는 대각 행렬의 곱으로 대각화 분해(Eigen decomposition)가 가능

고유값 가능 조건 - 일차독립

n*n 정방행렬 A가 고유값 분해가 가능하려면 행렬A가 n개의 일차독립인 고유벡터를 가져야 한다

일차독립이란 벡터들의 집합의 각 벡터들이 다른 벡터들의 일차 결합으로 표현될 수 없으면 일차독립이라고 한다

일차 결합이란 벡터에 상수를 곱하여 더한 것을 의미한다

$$\text{예를들어} v_1, v_2, v_3 \text{벡터가 존재할 때}$$

$$v_1 \neq av_2 + bv_3, v2 \neq av_1 + bv_3, v3 \neq av_1 + bv_2$$

$$\text{를 만족해야 한다}$$

고유값과 고유벡터의 계산

$$Av = \lambda v$$

$$Av - \lambda v = 0$$

$$(A - \lambda E)v = 0 \quad (\text{단}, v \neq 0, 0\text{은 영행렬})$$

$$\text{만약 } (A - \lambda E)\text{가 역행렬이 존재하면}$$

$$v = (A - \lambda E)^{-1} \cdot 0$$

$$v = 0\;(\text{v는 영행렬이 될 수 없으므로 모순이다})$$

$$\text{따라서 } A - \lambda E\text{는 역행렬이 존재하지 않음}$$

$$det(A - \lambda E) = 0\;(\text{det는 행렬식})$$

행렬을 하나 가정하여 고유값과 고유벡터를 구해보자

$$A = \begin{bmatrix} 2 & 0 & -2 \\ 1 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$$

$$det(A-\lambda E) = det(\begin{bmatrix} 2 & 0 & -2 \\ 1 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix})$$

$$= det(\begin{bmatrix} 2-\lambda & 0 & -2 \\ 1 & 1-\lambda & -2 \\ 0 & 0 & 1-\lambda \end{bmatrix})$$

$$= (2-\lambda)(1-\lambda)^2$$

$$\lambda = 1, 2 \ (1\text{은 중근})$$

λ에 대응하는 고유 벡터는 단일근이면 최대 1개, 이중근이면 최대 2개 ...n중근이면 최대 n개 의 고유벡터 존재

고유값을 이용하여 고유 벡터도 구해보자

$$\lambda = 2$$

$$\begin{bmatrix} 0 & 0 & -2 \\ 1 & -1 & -2 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$v_z = 0, v_x = v_y$$

$$\text{따라서 } v = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$$\lambda = 1$$

$$\begin{bmatrix} 1 & 0 & -2 \\ 1 & 0 & -2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$v_x = 2v_z$$

$$\text{따라서 } v = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 & -2 \\ 1 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1}$$

대칭행렬과 고유값 분해

대각원소를 중심으로 원소값들이 대칭되는 행렬

$$A^\top = A$$

실수로 구성된 원소의 대칭행렬은 항상 고유값 대각화가 가능하고 고유값 대각행렬은 직교행렬이다

$$A = PLP^{-1}$$

$$= PLP^\top$$

$$PP^\top = E$$

직교(orthogonal), 정규직교(orthonormal), 직교행렬(orthogmanal matrix)

직교와 정규직교

두 벡터가 수직이면 직교

$$v_1 \cdot v_2 = 0$$

어떤 행렬을 크기가 1인 단위 벡터로 만드는 것을 정규화라고 한다

$$v` = v/||v||$$

정규직교는 (직교+정규화)로 v1,v2가 단위벡터이면서 수직이면 정규직교라고 한다

$$\text{직교} = v_1 \cdot v_2 = 0$$

$$\text{정규직교} = v_1 \cdot v_2 = 0, ||v_1|| = 1, ||v_2|| = 1$$

직교행렬

$$A^{-1} = A^\top$$

$$AA^\top = E$$

직교행렬의 각 열벡터들은 서로 정규직교한 성질을 갖고 있다

즉, 직교행렬을 구성하는 열벡터들을 v1,v2...vn이라 했을 때 이들은 모두 단위벡터이면서 또한 서로 수직인 성질을 갖는다

SVD(특이값 분해)

SVD는 m*n행렬에 대해 행렬을 대각화 한다

$$A = U\Sigma V^\top$$

$$U : m * m \text{직교행렬} (AA^\top = U(\Sigma\Sigma^\top)V^\top)$$

$$V : n * n \text{직교행렬} (A^\top A = V(\Sigma^\top \Sigma)V^\top)$$

$$\Sigma : m * n \text{직사각대각행렬}$$

$$\text{U는 } AA^\top \text{를 고유값 분해하여 얻어진 직교 행렬}$$

U의 열벡터를 A의 left singular vector라고 부른다

$$\text{V는 } A^\top A \text{를 고유값 분해하여 얻어진 직교 행렬}$$

V의 열벡터를 A의 Right Singular Vector라고 부른다

$\Sigma\text{는} AA^\top, A^\top A\text{를 고유값 분해해서 나오는 고유값들의 square root를 대각원으로 하는 m*n 직사각 대각행렬, A의 특이값(Singular Vector) 라고 부른다}$

$$\Sigma = \begin{bmatrix} \sigma_1 & & \\ & \cdots & \\ & & \sigma_s \\ 0 & 0 & 0 \end{bmatrix} (m > n) \quad or \quad \Sigma = \begin{bmatrix} \sigma_1 & & & 0 \\ & \cdots & & 0 \\ & & \sigma_s & 0 \end{bmatrix} (m < n)$$

$AA^\top$와 $A^\top A$는 대칭행렬이므로 직교행렬로 고유값 분해가 가능

$AA^\top$와 $A^\top A$의 고유값들은 모두 0이상이며, 0이아닌 고유값은 모두 동일하다