9.4 특이값 분해

□ **정리 9.4.1**

A : $m \times n$ 행렬

(a) $\text{null}(A) = \text{null}(A^TA)$ (∵ A와 A^TA는 같은 영공간을 갖는다.)

(b) $\text{row}(A) = \text{row}(A^TA)$ (∵ A와 A^TA는 같은 행공간을 갖는다.)

(c) $\text{col}(A^T) = \text{col}(A^TA)$ (∵ A^T와 A^TA는 같은 열공간을 갖는다.)

(d) $\text{rank}(A) = \text{rank}(A^TA)$ (∵ A와 A^TA는 같은 랭크를 갖는다.)

[증명] (a)

(a) ① $\mathbf{x}_0$가 $A\mathbf{x} = \mathbf{0}$의 해라 하자. $\Rightarrow$ $\mathbf{x}_0$가 $A^TA\mathbf{x} = \mathbf{0}$의 해

∵

② $\mathbf{x}_0$가 $A^TA\mathbf{x} = \mathbf{0}$의 해라 하자. $\Rightarrow$ $\mathbf{x}_0$가 $A\mathbf{x} = \mathbf{0}$의 해

∵

(b)

(c)

(d)

□ **정리 9.4.2**

A : $m \times n$ 행렬

(a) A^TA는 직교대각화 가능하다.

(b) A^TA의 고유값이 음이 아니다.

[증명] (a) A^TA이 대칭행렬 $\therefore A^TA$는 직교대각화 가능

(b) A^TA는 직교대각화 가능

$\Rightarrow$ A^TA의 고유벡터로 구성된 R^n의 정규직교기저 $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ 존재.

대응 되는 고유값을 $\lambda_1, \lambda_2, \ldots, \lambda_n$ 이라 하면 $1 \le i \le n$에 대해

$$\|A\mathbf{v}_i\|^2 = A\mathbf{v}_i \cdot A\mathbf{v}_i = \mathbf{v}_i \cdot A^TA\mathbf{v}_i = \mathbf{v}_i \cdot \lambda_i \mathbf{v}_i = \lambda_i(\mathbf{v}_i \cdot \mathbf{v}_i) = \lambda_i \|\mathbf{v}_i\|^2 = \lambda_i$$

따라서 $\lambda_i \ge 0 \quad (1 \le i \le n)$

▣ **정의 1**

- A가 $m \times n$ 행렬이고 $\lambda_1, \lambda_2, \ldots, \lambda_n$가 A^TA의 고유값일 때 다음의 수를 A의 **특이값(singular value)**이라 한다.

$$\sigma_1 = \sqrt{\lambda_1},\ \sigma_2 = \sqrt{\lambda_2},\ \ldots,\ \sigma_n = \sqrt{\lambda_n}$$

예제 1 행렬 $\begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$의 특이값을 구하여라.

[풀이]

▣ **정의**

▪ A가 $m\times n$ 행렬일 때 A의 주대각(main diagonal)은
$m \geq n$ 일 때 원소 $a_{11}, a_{22}, \ldots, a_{nn}$이 있는 위치, $m \leq n$ 일 때 원소 $a_{11}, a_{22}, \ldots, a_{mm}$이 있는 위치.

▪ 이때 주대각에 있는 모든 원소를 **대각원소**라고 한다.

Ex

$$\begin{pmatrix} 1 & 2 \\ 3 & 0 \\ -1 & 4 \end{pmatrix} \qquad \begin{pmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{pmatrix}$$

□ **정리 9.4.4 <특이값 분해(SVD, singular value decomposition)>**

◆ A : 랭크가 k인 $m\times n$ 행렬

$$\Rightarrow \quad A = U\Sigma V^T = \left(\mathrm{u}_1 \quad \mathrm{u}_2 \quad \cdots \quad u_k \,|\, \mathrm{u}_{k+1} \cdots \mathrm{u}_m\right) \left(\begin{array}{cccc|c} \sigma_1 & 0 & \cdots & 0 & \\ 0 & \sigma_2 & \cdots & 0 & O_{k\times(n-k)} \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \cdots & \sigma_k & \\ \hline & O_{(m-k)\times k} & & & O_{(m-k)\times(n-k)} \end{array}\right) \begin{pmatrix} \mathrm{v}_1^T \\ \mathrm{v}_2^T \\ \vdots \\ \mathrm{v}_k^T \\ \hline \mathrm{v}_{k+1}^T \\ \vdots \\ \mathrm{v}_n^T \end{pmatrix}$$

여기서 U : $m\times m$ 직교행렬

V : $n\times n$ 직교행렬

Σ : 주 대각 성분이 A의 특이값이고 나머지 원소는 모두 0인 $m\times n$ 행렬

(a) $V = (\mathrm{v}_1\ \mathrm{v}_2 \cdots \mathrm{v}_n)$는 A^TA를 직교대각화한다.

(b) $\lambda_1, \lambda_2, \ldots, \lambda_k$가 V의 열벡터에 대응하는 A^TA의 영이 아닌 고유값이면,
Σ의 영이 아닌 대각성분은 $\sigma_1 = \sqrt{\lambda_1},\ \sigma_2 = \sqrt{\lambda_2},\ \ldots, \sigma_k = \sqrt{\lambda_k}$ (: 특이값)

(c) V의 열벡터는 $\sigma_1 \geq \sigma_2 \geq \cdots \sigma_k > 0$을 만족시키도록 배열되어 있다.

(d) $\mathrm{u}_i = \dfrac{1}{\|A\mathrm{v}_i\|}A\mathrm{v}_i = \dfrac{1}{\sigma_i}A\mathrm{v}_i, \quad i = 1, 2, \ldots, k$

(e) $\{\mathrm{u}_1, \mathrm{u}_2, \cdots, \mathrm{u}_k\}$는 $\mathrm{col(A)}$의 정규직교기저이다.

(f) $\{\mathrm{u}_1, \mathrm{u}_2, \cdots, \mathrm{u}_k, \mathrm{u}_{k+1}, \ldots, \mathrm{u}_m\}$은 $\{\mathrm{u}_1, \mathrm{u}_2, \cdots, \mathrm{u}_k\}$로부터 R^m의 정규직교기저로 확장한 것이다.

[증명] (A가 랭크가 k인 $n \times n$ 행렬인 경우에)

A^TA는 대칭행렬 $\Rightarrow$ 직교대각화 가능(고유값 분해) : $A^TA = VDV^T$

이때 $V = (\mathrm{v}_1\ \mathrm{v}_2\ \cdots\ \mathrm{v}_n)$의 열벡터는 A^TA의 정규직교 고유벡터

D의 대각성분 $\lambda_1, \lambda_2, \ldots, \lambda_n$은 V의 열벡터에 대응하는 A^TA의 고유값

$\mathrm{rank}(A) = k \quad \Rightarrow \quad \mathrm{rank}(A^TA) = k$

$\Rightarrow \quad \mathrm{rank}(D) = k \quad$ ($\because$ D와 A^TA가 닮음)

$$D = \begin{pmatrix} \lambda_1 & & & & & 0 \\ & \lambda_2 & & & & \\ & & \ddots & & & \\ & & & \lambda_k & & \\ & & & & 0 & \\ & & & & & \ddots \\ 0 & & & & & 0 \end{pmatrix}$$ 여기서 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k > 0$ 이다.

$\{A\mathrm{v}_1, A\mathrm{v}_2, \cdots, A\mathrm{v}_n\}$ (A의 상(image)의 집합) 은 직교집합이다.

($\because$ $i \neq j$ 때 v_i 와 v_j가 직교 $\Rightarrow$ $A\mathrm{v}_i \cdot A\mathrm{v}_j = \mathrm{v}_i \cdot A^TA\mathrm{v}_j = \mathrm{v}_i \cdot (\lambda_j \mathrm{v}_j) = \lambda_j(\mathrm{v}_i \cdot \mathrm{v}_j) = 0$)

D에서 처음 k개의 대각원소들은 영이 아니고, 또한 $\| A\mathrm{v}_i \|^2 = \lambda_i$ (정리 9.5.2(b) 증명에서)이므로

$\{A\mathrm{v}_1, A\mathrm{v}_2, \cdots, A\mathrm{v}_n\}$에서 처음 k개의 벡터들을 영벡터가 아니다.

$S = \{A\mathrm{v}_1, A\mathrm{v}_2, \cdots, A\mathrm{v}_k\}$는 영이 아닌 벡터들로 이루어진 A의 열공간의 직교집합이다.

A의 열공간의 차원은 $\mathrm{rank}(A) = \mathrm{rank}(A^TA) = k$ 이므로 S는 $\mathrm{col}(A)$의 기저이다.

$\Rightarrow$ 정규화 $\mathrm{u}_i = \dfrac{1}{\| A\mathrm{v}_i \|} A\mathrm{v}_i = \dfrac{1}{\sqrt{\lambda_i}} A\mathrm{v}_i = \dfrac{1}{\sigma_i} A\mathrm{v}_i$ ($\Leftrightarrow$ $A\mathrm{v}_i = \sqrt{\lambda_i}\,\mathrm{u}_i = \sigma_i \mathrm{u}_i$)

따라서 $\mathrm{col(A)}$의 정규직교기저 $\{\mathrm{u}_1, \mathrm{u}_2, \cdots, \mathrm{u}_k\}$를 얻는다.

이 기저에서 R^m의 정규직교기저로 확장한다. $\{\mathrm{u}_1, \mathrm{u}_2, \cdots, \mathrm{u}_k, \mathrm{u}_{k+1}, \ldots, \mathrm{u}_m\}$

$U = (\mathrm{u}_1\ \mathrm{u}_2\ \cdots\ \mathrm{u}_k\ \mathrm{u}_{k+1}\ \cdots\ \mathrm{u}_m)$라 하면 $\Rightarrow$ U는 직교행렬

$$\Sigma = \begin{pmatrix} \sigma_1 & & & & & 0 \\ & \sigma_2 & & & & \\ & & \ddots & & & \\ & & & \sigma_k & & \\ & & & & 0 & \\ & & & & & \ddots \\ 0 & & & & & 0 \end{pmatrix}$$ 라 하면,

$$U\Sigma = (\sigma_1 \mathrm{u}_1\ \sigma_2 \mathrm{u}_2\ \cdots\ \sigma_k \mathrm{u}_k\ 0\ \cdots\ 0) = (A\mathrm{v}_1\ A\mathrm{v}_2\ \cdots\ A\mathrm{v}_k\ A\mathrm{v}_{k+1}\ \cdots\ A\mathrm{v}_n) = AV$$

$\boxed{\text{예제 2}}$ $A=\begin{pmatrix}1 & 1\\ 0 & 1\\ 1 & 0\end{pmatrix}$의 특이값 분해를 찾아라.

[풀이]

9.5 특이값 분해를 이용한 자료압축

[Note]

- 특이값 분해에서 행렬의 영행과 영열은 불필요하므로 다음과 같은 형태의 분해를 얻을 수 있다.

$$A = \begin{pmatrix} \mathrm{u}_1 & \mathrm{u}_2 & \cdots & \mathrm{u}_k \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_k \end{pmatrix} \begin{pmatrix} \mathrm{v}_1^T \\ \mathrm{v}_2^T \\ \vdots \\ \mathrm{v}_k^T \end{pmatrix}$$

: **A의 축소된 특이값 분해**--(∗)

$$A = U_1 \Sigma_1 V_1^T$$

이때 U_1은 $m \times k$ 행렬, Σ_1은 $k \times k$ 가역행렬, V_1^T은 $k \times n$행렬

- (∗) 식을 전개하면 **A의 축소된 특이값 전개**를 얻는다.

$$A = \sigma_1 \mathrm{u}_1 \mathrm{v}_1^T + \sigma_2 \mathrm{u}_2 \mathrm{v}_2^T + \cdots + \sigma_k \mathrm{u}_k \mathrm{v}_k^T$$

예제 1 행렬 $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$의 (축소된 특이값 분해와) 축소된 특이값 전개를 구하여라.

즉, $A = \sigma_1 \mathrm{u}_1 \mathrm{v}_1^T + \sigma_2 \mathrm{u}_2 \mathrm{v}_2^T$로 나타내어라.

[풀이]