· 정방 행렬은 고윳값 분해를 이용하여 고윳값과 고유벡터를 찾을 수 있었다. 정방행렬이 아닌경우 특이분해를 이용할 수 있다.

· 특이값 분해 (singular value)

- N×M 크기의 행렬A를 다음과 같은 3개의 곱으로 나타내는 것을 특이분해라고 한다.

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

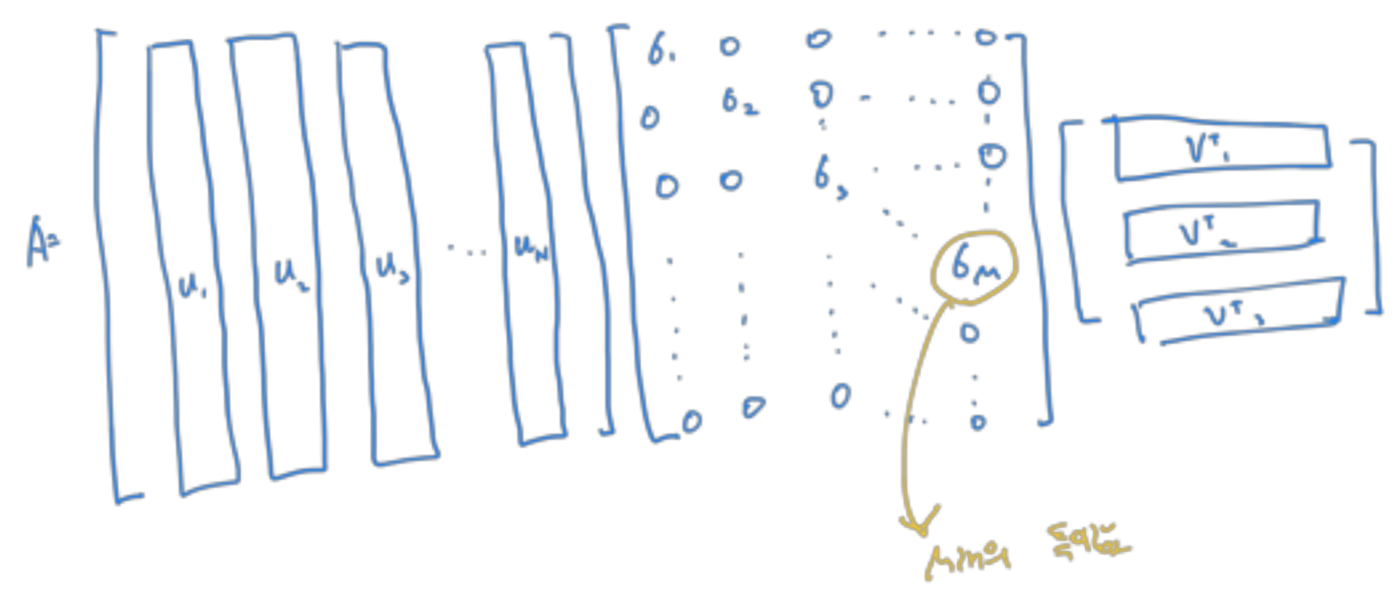
특이값 ← $\Sigma \in R^{N\times M}$: 대각성분이 양수인 대각행렬. 큰수부터 작은수로 배열한다.

왼쪽특이벡터 ← $U \in R^{N\times N}$: N×N차원 정방행렬. 모든 열벡터가 단위벡터이고, 서로 직교한다.

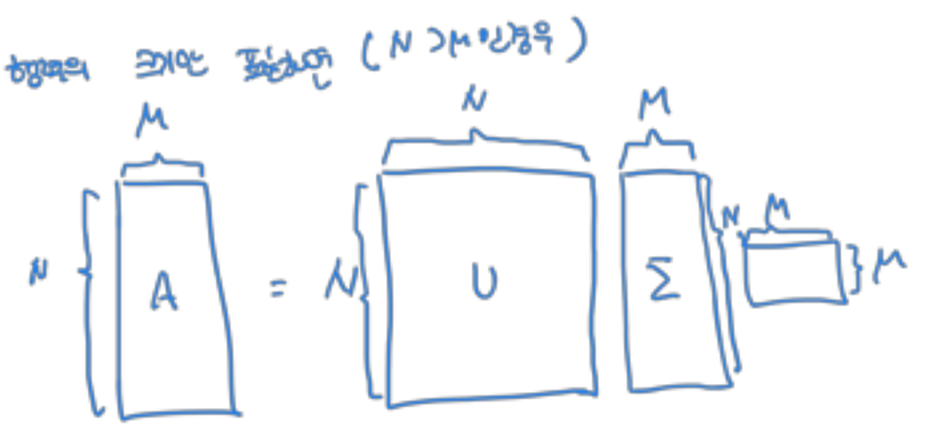
오른쪽특이벡터 ← $V \in R^{M\times M}$: M×M차원 정방행렬. 모든 열벡터가 단위벡터이고, 서로 직교한다.
↳ 길이가 1인 벡터

∘ 특이값의 갯수는 행렬의 열과 행의 갯수중 작은 값과 같다

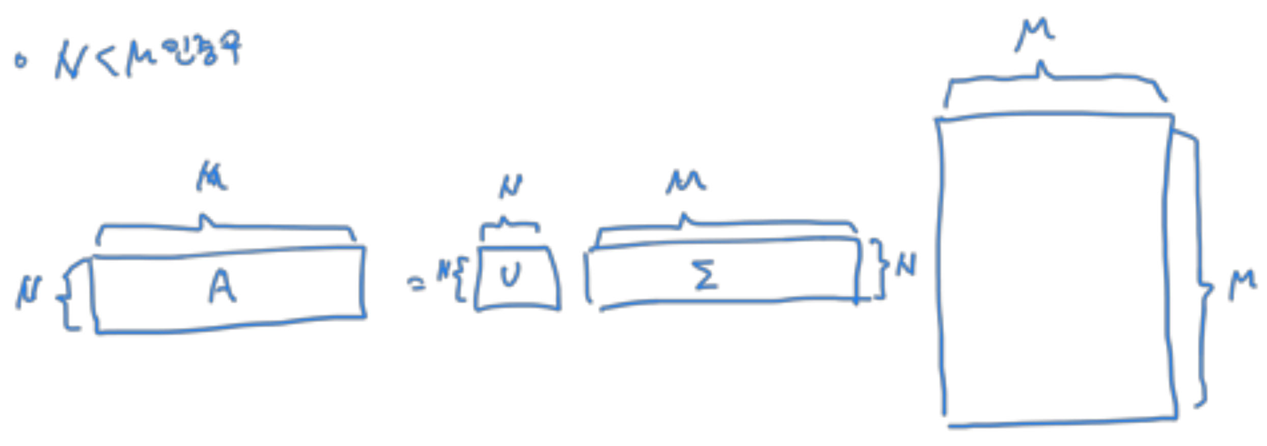
∘ 만약 N>M 인 경우,



∘ 행렬의 크기만 표시하면 (N>M인경우)



∘ N<M 인경우



∘ 특이값과 특이벡터의 관계

- 행렬 V는 정규직교 행렬이므로 전치행렬이 역행렬이다.

$$V^T = V^{-1}$$

- 특이분해된 등식의 양변에 V를 곱하면,

$$AV = U\Sigma V^T V = U\Sigma \Rightarrow A[v_1\ v_2 \cdots v_M] = [u_1\ u_2 \cdots u_N]\begin{bmatrix}\sigma_1 & 0 & \cdots & \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & & \ddots & \end{bmatrix}$$

- 행렬 A를 곱하여 정리하면, M이 N보다 클때는

$$[Av_1, Av_2, Av_3 \cdots Av_M] = [u_1\sigma_1, u_2\sigma_2 \cdots u_N\sigma_N]$$

- 반대로, N이 M보다 큰경우는...

$$[Av_1, Av_2, Av_3 \cdots Av_M] = [u_1\sigma_1, u_2\sigma_2 \cdots u_M\sigma_M]$$

즉, i번째 특이값 σ_i와 특이벡터 u_i, v_i은 다음과 같은 관계가 있다.

$$Av_i = \sigma_i u_i \quad (i = 1 \dots \min(M,N))$$

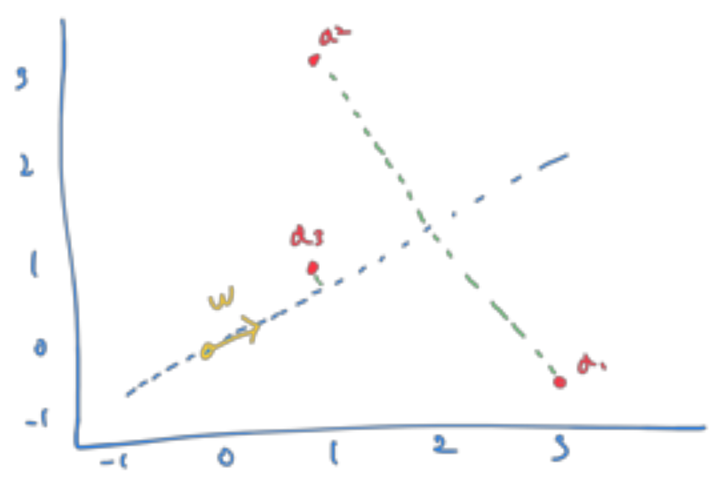
∘ 특이분해와 고유분해의 관계

- 행렬 A의 분산행렬 A^TA는

$$A^TA = (V\Sigma^T U^T)(U\Sigma V^T) = V\Lambda V^T$$

· 1차원 근사

- 2차원 평면위에 3개의 2차원 벡터 a_1, a_2, a_3 가 있다. 원점을 지나면서 모든점들과 가능한 가까이 있는 직선을 만들기 위해



- 벡터 w와 점 a_i의 거리의 제곱은 다음처럼 계산할 수 있다.

$$\|a_i^{\perp w}\|^2 = \|a_i\|^2 - \|a_i^{\| w}\|^2 = \|a_i\|^2 - (a_i^T w)^2$$

예를들어 벡터 a_1, a_2, a_3를 행벡터로 가지는 행렬 A를 가정하면,

$$A = \begin{bmatrix} a_1^T \\ a_2^T \\ a_3^T \end{bmatrix}$$

행벡터의 놈의 제곱의 합은 행렬의 놈이므로 모든 점들과의 거리의 제곱의 합은 행렬의 놈으로 계산된다.

$$\sum_{i=1}^{3} \|a_i^{\perp w}\|^2 = \sum_{i=1}^{3} \|a_i\|^2 - \sum_{i=1}^{3} (a_i^T w)^2 = \|A\|^2 - \|Aw\|^2$$

$$\arg\max_w \|Aw\|^2$$

· 랭크-1 근사문제

a_i의 투영성분벡터는

$$a_i^{\| w} = (a_i^T w)w$$ 이므로 w벡터를 이용하면,

N개의 M차원 벡터 $a_1, a_2, a_3 \cdots a_N$ ($a_i \in R^M$)를 1차원으로 투영하여 가장 비슷한 N개의 1차원 벡터 $a_1^{\| w}, a_2^{\| w}, \dots, a_N^{\| w}$ ($a_i^{\| w} \in R^1$)를 만들 수 있다.

$$A' = \begin{bmatrix} (a_1^{\| w})^T \\ (a_2^{\| w})^T \\ \vdots \\ (a_N^{\| w})^T \end{bmatrix} = \begin{bmatrix} a_1^T ww^T \\ a_2^T ww^T \\ \vdots \\ a_N^T ww^T \end{bmatrix}$$

$$= \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_N^T \end{bmatrix} ww^T = Aww^T$$

이 답은 원래 행렬 A에 랭크-1 행렬 ww^T를 곱해서 원래의 행렬 A와 가장 비슷한 행렬 A'을 만드는 문제와 같다.
최소

$$\arg\min_w \|A - A'\| = \arg\min_w \|A - Aww^T\|$$

랭크-1 근사문제 (rank-1 approximation problem)

∘ 랭크-K 근사문제

$$\arg\min_{w_1 \cdots w_k} \|A - Aww^T\|$$

랭크-K 근사문제 (Rank-K approximation problem)

특이값 분해