Agenda:

- HWs due Saturday nights now
- HW10 selfgrade still due Friday
- Questions about courses? Reach out!
- SVD!
- Worksheet

Motivations:

- helps us solve for min norm solution
- decomposes non-square matrix
  - generalizes Spectral Theorem applications to non-symmetric/square matrices
- many more!

Singular Value Decomposition

Given $A \in \mathbb{R}^{m \times n}$:

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

where

- U : orthonormal matrix $\in \mathbb{R}^{m \times m}$
- Σ : non-square diagonal matrix w/ "singular values" along diagonal $\sigma_i = \sqrt{\lambda_i}$ where λ_i is eigenvalue of (AA^T) or (A^TA)
- V^T: orthonormal matrix $\in \mathbb{R}^{n \times n}$

① Full SVD: $A = U\Sigma V^T$

$$A = \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_m \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_m & \end{bmatrix} \begin{bmatrix} \vec{v}_1 \\ \vdots \\ \vec{v}_n \end{bmatrix}$$

(assuming A is wide matrix)

② Compact SVD: $A = U_r \Sigma_r V^T$ where $\text{Rank}(A) = r$

$$A = \left[\begin{array}{ccc|ccc} \vec{u}_1 & \cdots & \vec{u}_r & \vec{u}_{r+1} & \cdots & \vec{u}_m \end{array}\right] \left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_r & \\ \hline & 0 & & 0 \end{array}\right] \left[\begin{array}{c} \vec{v}_1^T \\ \vdots \\ \vec{v}_r^T \\ \hline \vec{v}_{r+1}^T \\ \vdots \\ \vec{v}_n \end{array}\right]$$

$$= \left[\begin{array}{c|c} U_r & U_{m-r} \end{array}\right] \left[\begin{array}{c|c} \Sigma_r & 0_{r \times (n-r)} \\ \hline 0_{(m-r) \times r} & 0_{(m-r) \times (n-r)} \end{array}\right] \left[\begin{array}{c} V_r^T \\ \hline V_{n-r}^T \end{array}\right]$$

$$= \left[\begin{array}{c|c} U_r & U_{m-r} \end{array}\right] \begin{bmatrix} \Sigma_r V_r^T + 0 V_{n-r}^T \\ 0 V_r^T + 0 \cdot V_{n-r}^T \end{bmatrix}$$

$$= \left[\begin{array}{c|c} U_r & U_{m-r} \end{array}\right] \begin{bmatrix} \Sigma_r V_r^T \\ 0 \end{bmatrix}$$

$$= U_r \Sigma_r V_r^T + U_{m-r} \cdot 0$$

$$A = U_r \Sigma_r V_r^T$$

☆ For proof of SVD, refer to lecture and Note 14

③ Outer-Product Form: $A = \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^T$

$$A = \underset{(m \times m)}{\begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_m \end{bmatrix}} \underset{(m \times n)}{\left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_m & \end{array}\right]} \underset{(n \times n)}{\begin{bmatrix} \vec{v}_1^T \\ \vdots \\ \vec{v}_n^T \end{bmatrix}}$$

$$= \underset{(m \times m)}{\begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_m \end{bmatrix}} \underset{(m \times n)}{\begin{bmatrix} \sigma_1 v_1^T \\ \vdots \\ \sigma_m v_m^T \end{bmatrix}}$$

$$= \vec{u}_1(\sigma_1 \vec{v}_1^T) + \vec{u}_2(\sigma_2 \vec{v}_2^T) + \cdots + \vec{u}_m \sigma_m \vec{v}_m^T$$

$$= \sigma_1(\vec{u}_1 \vec{v}_1^T) + \cdots + \sigma_m \vec{u}_m \vec{v}_m^T$$

if $\text{rank}(A) = r \Rightarrow \sigma_{r+1} \ldots \sigma_m = 0$

$$= \sigma_1(\vec{u}_1 \vec{v}_1^T) + \cdots + \sigma_r \vec{u}_r \vec{v}_r^T$$

$$A = \sum_{i=0}^{r} \sigma_i \vec{u}_i \vec{v}_i^T \quad \checkmark$$

Outer Product:

$\vec{u}_i \vec{v}_i^T$

$$= \underset{(m \times 1)}{\begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}} \underset{(1 \times n)}{\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}}$$

$$= \underset{(m \times n)}{\begin{bmatrix} u_1 v_1 & \cdots & u_1 v_n \\ \vdots & \ddots & \vdots \\ u_m v_1 & \cdots & u_m v_n \end{bmatrix}}$$

Connecting Spectral Theorem to SVD

- In Spectral Theorem, $A \in \mathbb{R}^{n \times n}$ is symmetric /square where $\text{rank}(A) = r$

  $A = V \Lambda V^T$

  where Λ is square & diagonal w/ eigenvalues of A

  $V = \begin{bmatrix} V_r & V_{n-r} \end{bmatrix}$

  - V_r spans $\text{Col}(A) = \text{Col}(A^T)$
  - V_{n-r} spans $\text{Null}(A) = \text{Null}(A^T)$

  $\}$ why? b/c symmetric!

- In SVD, $A \in \mathbb{R}^{m \times n}$ not symmetric/square, $\text{rank}(A) = r$:

  $A = U \Sigma V^T$

  where Σ: non-square diagonal matrix of singular values

  $U = \begin{bmatrix} U_r & U_{m-r} \end{bmatrix} \in \mathbb{R}^{m \times m}$, orthonormal
  - ☆ U_r spans $\text{Col}(A)$
  - U_{m-r} orthogonal to $\text{Col}(A)$, spans $\text{Null}(A^T)$

  $V = \begin{bmatrix} V_r & V_{n-r} \end{bmatrix} \in \mathbb{R}^{n \times n}$, orthonormal
  - V_r spans $\text{Col}(A^T)$
  - ☆ V_{n-r} orthogonal to $\text{Col}(A^T)$, spans $\text{Null}(A)$

  for a symmetric matrix these are the same! $U = V$ b/c $A = A^T$