Why is SVD helpful?

① It helps give insight ~~into~~ other ML methods.

② It can be useful for dimension reduction.

<u>Dimension Reduction</u>

$$X = UDV^T$$
$$= \sum_{j=1}^{p} \delta_j u_j v_j^T \quad \text{w/}$$

$$\delta_1 \geq \delta_2 \geq \delta_3 \geq \cdots \geq \delta_p \geq 0$$

<u>Idea</u>: If at somepoint k the singular values $\{\delta_j\}_{j \geq k}$ are very small, their relative contribution to the weighted sum is very small:

$$X = \delta_1 u_1 v_1^T + \delta_2 u_2 v_2^T + \cdots + \delta_k u_k v_k^T + \delta_{k+1} u_{k+1} v_{k+1}^T + \cdots + \delta_p u_p v_p^T$$

If they aren't adding much, who needs them?

If I let $\{\delta_j\}_{j>k}$ all equal 0, we get:

$$X \approx \delta_1 u_1 v_1^T + \cdots + \delta_k u_k v_k^T \ (+0 + \cdots + 0)$$

$$= \underbrace{U_k}_{n \times k} \ \underbrace{D_k}_{k \times k} \ \underbrace{V_k^T}_{k \times p}$$

$$= \begin{bmatrix} | & & | \\ u_1 & \cdots & u_k \\ | & & | \end{bmatrix} \begin{bmatrix} \delta_1 & & 0 \\ & \ddots & \\ 0 & & \delta_k \end{bmatrix} \begin{bmatrix} — v_1^T — \\ \vdots \\ — v_k^T — \end{bmatrix}$$

Low Rank Approximation of X.

Instead of storing $n \times p$ elements of the matrix X, I now only have to store:

$$n \times k + k + k \times p =$$

$$(n+p+1) \times k$$