Singular Value Decomposition

$X = USV$ where X is $n \times p$ and

- U is an $n \times n$ orthogonal matrix. $U^t U = I$.
- V^t is an $p \times p$ orthogonal matrix. $V^t V = I$.
- S is an $n \times p$ "diagonal" matrix $\begin{pmatrix} \sigma_1 & & 0 \\ & \sigma_2 & \\ 0 & & \ddots \; \sigma_p \\ \hline & 0 & \end{pmatrix}$

The non-zero entries of S are called the singular values of S.

Relationship with PCA

Recall that the PCA decomposition can be expressed as:

$$X^t X = E^t D E$$

Suppose that X is decomposed as $X = USV^t$

$$\begin{aligned} X^t X &= V S^t U^t U S V \\ &= V S^t S V^t \end{aligned}$$

$$S^t S = \left(\begin{array}{ccc|c} \sigma_1 & & 0 & \\ & \sigma_2 & & 0 \\ 0 & & \ddots \; \sigma_p & \end{array} \right) \begin{pmatrix} \sigma_1 & & 0 \\ & \sigma_2 & \\ 0 & & \ddots \; \sigma_p \\ \hline & 0 & \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \sigma_2^2 & \\ 0 & & \ddots \; \sigma_p^2 \end{pmatrix}$$

So $S^t S$ is a diagonal matrix.

PCA	SVD
$X^t X = E^t D E$	$X^t X = V S^t S V^t$

$\Longrightarrow E = V^t$ and $D = S^t S$

So: SVD immediately gives PCA!

Why SVD

Suppose X contains image data.

- 500 images
- Each image is $200 \times 200 \Longrightarrow 40{,}000$ pixels.

$\Longrightarrow X$ is a $500 \times 40{,}000$

This is called the $p >> n$ situation.

PCA needs to compute the matrix $X^t X \longleftarrow 40{,}000 \times 40{,}000$!

$X^t X$ has $800{,}000{,}000 = 8 \times 10^8$ entries.

1 floating point # is 64 bits
is 8 bytes

So $X^t X$ is $8 \times 8 \times 10^8 = 64 \times 10^8$ bytes
$= 64$ GB of data.

SVD can calculate the <u>same</u> eigenvalues without calculaing $X^t X$.
Much more efficient when $p >> n$.