Assume $n \geq p$.

Theorem 1. *For any $X \in \mathbb{R}^{n \times p}$, there exists SVD decomposition:*

$$X = P \Lambda Q^T \tag{1}$$

where $P \in \mathbb{R}^{n \times n}$, $Q \in \mathbb{R}^{p \times p}$ are orthogonal matrices, and $\Lambda = diag\left(\sigma_1, \sigma_2, \cdots, \sigma_p\right) \in \mathbb{R}^{n \times p}$ is a rectangular diagonal matrix of nonnegative entries $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0$.

Proof. We will prove for X full rank[1]. Proof is in three (rather simple) parts.

1. Show that $X^T X$ is a positive definite[2] matrix.
2. Eigendecompose $X^T X$.
3. Construct SVD decomposition.

Part 1 $X^T X$ is obviously symmetric. To show positive definiteness, note that for any $v \in \mathbb{R}^p$,

$$v^T \left(X^T X\right) v = (Xv)^T (Xv) = |Xv|^2 \geq 0$$

So $X^T X$ is a symmetric positive semi-definite matrix.

Part 2 Since $X^T X$ is a symmetric positive-semidefinite matrix, by spectral theorem, the following eigendecomposition exists:

$$X^T X = Q D Q^T = Q \tilde{\Lambda}^2 Q^T \tag{2}$$

where $Q \in \mathbb{R}^{p \times p}$ is orthogonal, and $D \in \mathbb{R}^{p \times p}$ is a diagonal matrix of nonnegative entries. Without loss of generality $D_{11} \geq D_{22} \geq \cdots \geq D_{pp} \underbrace{> 0}_{X \text{ full rank}}$

Then, we introduce $\tilde{\Lambda} = \operatorname{diag}\left(\sqrt{D_{11}}, \cdots, \sqrt{D_{pp}}\right)$, then $\tilde{\Lambda}^2 = D$.

Part 3 Define $\tilde{P} = X Q \tilde{\Lambda}^{-1} \in \mathbb{R}^{n \times p}$. Then $\tilde{P}^T \tilde{P} = \tilde{\Lambda}^{-1} Q^T X^T X Q \tilde{\Lambda}^{-1} = I_p$, so $\tilde{P}$ is a rectangular matrix with orthonormal columns. Then we have:

$$X = \left(X Q \tilde{\Lambda}^{-1}\right) \tilde{\Lambda} Q^T = \underbrace{\tilde{P}}_{n \times p} \underbrace{\tilde{\Lambda}}_{p \times p} \underbrace{Q^T}_{p \times p} \tag{3}$$

<u>This is very similar to what we want to show!</u> To get the form in the lecture note, define $P = \left[\tilde{P} \, \tilde{P}_\perp\right] \in \mathbb{R}^{n \times n}$ where $\tilde{P}_\perp$ is orthogonal complement to $\tilde{P}$.

[1] Meaning, X has rank p.

[2] Only nonnegative eigenvalues

Also define $\Lambda = \begin{bmatrix} \tilde{\Lambda} \\ 0 \end{bmatrix} \in \mathbb{R}^{n \times p}$. Now we have:

$$X = \tilde{P}\tilde{\Lambda}Q^T = P\Lambda Q^T \tag{4}$$

$\square$

Example 1. Here is a trivial SVD decomposition:

$$\underbrace{\begin{pmatrix} -2 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}}_{X} = \underbrace{\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{P} \underbrace{\begin{pmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}}_{\Lambda} \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{Q^T} \tag{5}$$

Remark 1. SVD (1) can be rewritten as:

$$X = \sum_{i=1}^{p} \sigma_i p_i q_i^T \tag{6}$$

where p_i is the i^{th} column vector of P and q_i^T is the i^{th} row vector of Q^T. Note that ${\sigma_i}$[3] are ordered from greatest to least. Also p_i and q_i have unit norm.

To find rank-1 approximation of X, take

$$X \approx \sigma_1 p_1 q_1^T = \tilde{X} \tag{7}$$

To find rank-2 approximation of X, take

$$X \approx \sigma_1 p_1 q_1^T + \sigma_2 p_2 q_2^T = \tilde{X} \tag{8}$$

[3] singular values