MTH 100 : Lecture 38

Singular Value Decomposition

Let A be an $m \times n$ matrix.

Then $(A^TA)^T = A^T(A^T)^T = A^TA$

Therefore A^TA is a symmetric $n \times n$ matrix.

and can be orthogonally diagonalized.

Let $\{v_1, v_2, \dots, v_n\}$ be an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of A^TA with corresponding eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$.

Then $\|Av_i\|^2 = \langle Av_i, Av_i \rangle$

$$= (Av_i)^T(Av_i) = (v_i^T A^T)(Av_i)$$

$$= v_i^T(A^TA)v_i$$

$$= v_i^T(\lambda_i v_i) = \lambda_i v_i^T v_i$$

$$= \lambda_i \|v_i\|^2$$

$$= \lambda_i \cdot 1 = \lambda_i$$

So, $\|Av_i\|^2 = \lambda_i \Rightarrow \|Av_i\| = \sqrt{\lambda_i}$ for $i = 1, 2, \dots, n$

Thus $\lambda_i \geq 0$ for $i = 1, 2, \dots, n$

Therefore all the eigenvalues of the matrix $A^T A$ are nonnegative.

Definition:

Let A be an $m \times n$ matrix.

The singular values of A are the square roots of the eigen values of $A^T A$ denoted by $\sigma_1, \sigma_2, \dots, \sigma_n$ arranged in descending order

i.e. $\sigma_i = \sqrt{\lambda_i}$ for $i = 1, 2, \dots, n$

$$\sigma_1 \geq \sigma_2 \geq \dots\dots\dots \geq \sigma_n$$

Note that the singular values are the lengths of the vectors

$$Av_1, Av_2, \dots\dots, Av_n.$$

Proposition: Suppose $\{v_1, \ldots, v_n\}$ is an **orthonormal basis** for $\mathbb{R}^n$ consisting of eigenvectors of A^TA with corresponding eigenvalues arranged so that

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$$

Suppose that A has r nonzero singular values.

Then $\{Av_1, Av_2, \ldots, Av_r\}$ is an orthogonal basis for $\operatorname{Col} A$.

and $\operatorname{rank} A = r$

Proof: First note that for $j > r$, $\|Av_j\| = \sqrt{\lambda_j} = \sigma_j = 0$

$\Rightarrow Av_j = 0$

Now for $i, j \leq r$ $(i \neq j)$

we have $\langle Av_i, Av_j \rangle$

$$= (Av_i)^T (Av_j)$$

$$= (v_i^T A^T)(Av_j)$$

$$= v_i^T (A^T A) v_j$$
$$= v_i^T \lambda_j v_j$$
$$= \lambda_j v_i^T v_j$$
$$= \lambda_j \langle v_i, v_j \rangle = 0$$

(Since $\{v_1, v_2, \dots, v_n\}$ is an orthonormal basis for $\mathbb{R}^n$)

So, $\{Av_1, Av_2, \dots, Av_r\}$ is an orthogonal set of nonzero vectors and are therefore linearly independent.

Next observe that the vectors $Av_1, Av_2, \dots Av_r, \dots Av_n$ belong to Col A (of course $Av_{r+1} = \dots = Av_n = 0$)

Now let $y \in \text{Col } A$

Then $y = Ax$ for some $x \in \mathbb{R}^n$

Since $\{v_1, v_2, \dots, v_n\}$ is a basis for $\mathbb{R}^n$, x can be expressed as

$$x = c_1 v_1 + c_2 v_2 + \dots + c_n v_n$$

Note that if $A = [a_1\ a_2 \dots a_n]$ then

$$Ax = [a_1\ a_2 \dots a_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$= x_1 a_1 + x_2 a_2 + \dots + x_n a_n \in \text{Col} A$$

Then $y = Ax = A(c_1 v_1 + c_2 v_2 + \dots + c_n v_n)$

$$\Rightarrow y = c_1 A v_1 + c_2 A v_2 + \cdots + c_r A v_r + c_{r+1} A v_{r+1} + \cdots + c_n A v_n$$

$$= c_1 A v_1 + c_2 A v_2 + \cdots + c_r A v_r$$

Thus the vectors $Av_1, Av_2, \ldots, Av_r$ span Col A

Therefore the set of vectors $\{Av_1, Av_2, \ldots, Av_r\}$ forms an orthogonal basis for col A

and $\text{Rank}\, A = \dim(\text{col}\, A) = r$

Singular Value Decomposition (SVD)

Theorem (Singular Value Decomposition of a matrix):

Let A be an $m \times n$ matrix with rank r.

Then A can be factored as a product

$$A = U \Sigma V^T \quad \text{as follows:}$$

- Σ is an $m \times n$ matrix containing an $r \times r$ diagonal matrix D with the r non-zero singular values of A, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, along the main diagonal. D is placed in the upper left corner of Σ. Remaining entries of Σ are zero.

- U is an $m \times m$ orthogonal matrix and V is an $n \times n$ orthogonal matrix.

- The matrix V has as its columns the orthonormal basis $\{v_1, v_2, \dots, v_n\}$ of eigenvectors of A^TA.

- In order to obtain U, we take r vectors Av_i corresponding to the non-zero singular values, extend to an orthogonal basis of $\mathbb{R}^m$ using the Gram-Schmidt Process (This step is necessary only in case $r < m$) and finally normalize the vectors to obtain an orthonormal basis $\{u_1, u_2, \dots, u_m\}$. U has the vectors u_i as its columns.

Note: Any factorization $A = U \Sigma V^T$, with U and V as orthogonal matrices, Σ as described above is called a Singular Value Decomposition of SVD of A.
Note that U and V are not uniquely determined by A, but the diagonal entries of Σ are necessarily the singular values of A.