2.8. Singular Value Decomposition

Notes are created by Busra Tugce Gurbuz based on *Deep Learning* book (Goodfellow et al.)

- The **singular value decomposition** (SVD) provides another way to factorize a matrix, into singular vectors and singular values. SVD allows us to discover some of the same kind of information as the eigendecomposition.
- However, SVD is more generally applicable. Every real matrix has a SVD but the same is not true for eigendecomposition.
    - For example, if a matrix is not square, the eigendecomposition is not defined, and we must use a SVD instead.
- SVD is similar to eigendecomposition, but this time we will write it as a product of three matrices:

orthogonal matrix — diagonal matrix, not square — orthogonal matrix

$$A = U D V^T$$

A: m×n matrix; U: m×m matrix; D: m×n matrix; V^T: n×n matrix

$$U^TU = UU^T = I$$
$$V^TV = VV^T = I$$

D: diagonal matrix of singular values of matrix A

U: columns of U are left singular vectors

V: columns of V are right singular vectors

$$A = \begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_m \\ | & | & & | \end{bmatrix} \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \\ \hline & & 0 & \end{bmatrix} \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{bmatrix}^T$$

- We can actually interpret the SVD of A in terms of the eigendecomposition of functions of A. The left singular vectors of A are the eigenvectors of AA^T. The right singular vectors of A are the eigenvectors of A^TA. The non-zero singular values of A are the square roots of the eigenvalues of A^TA. The same is true for AA^T.

Proof for right singular vectors:

eigendecomposition of A^TA:

$$A^TA = [Q\Lambda Q^T]^T[Q\Lambda Q^T]$$
$$= Q\Lambda^T Q^T Q\Lambda Q^T$$
$$= Q\Lambda^T I \Lambda Q^T = Q\Lambda^T\Lambda Q^T$$

right singular vector of A: U

$$UDV^T = Q\Lambda^T\Lambda Q^T$$

$$Q = U \qquad Q^T = V^T$$
$$\Lambda^T\Lambda = D$$

- SVD is the first step in ML for dimensionality reduction.
    - It is also thought as data-driven generalization of the Fourier transform (FFT).
    - It is also the basis of PCA.