12 Linear Algebra 2 Lecture 12: Singular value decomposition (05/06/2020)

12.1 Singular values

Consider the following matrix transformation A applied to a unit circle.

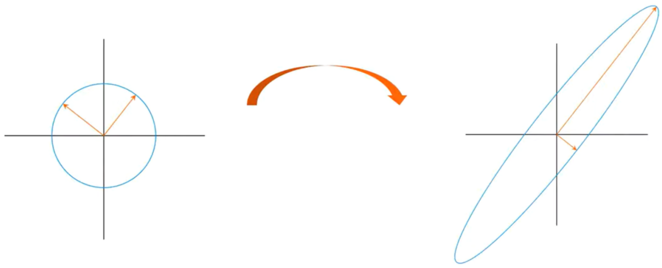

Figure 1: How the unit circle changes to an ellipse under the linear transformation of matrix A

Where the matrix A is given as:

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 2 \end{bmatrix} \tag{1}$$

The semi-major axis for the created ellipse will be $|\lambda_{max}| = \frac{1}{2}(3+\sqrt{17})$. The semi-minor axis will then be given as: $|\lambda_{min}| = \frac{1}{2}(\sqrt{17}-3)$. Let's now consider a more tricky situation. We start with a unit sphere in $\mathbb{R}^3$. We then map this sphere to an ellipse on a 2 diminsional plane. The matrix A for this transformation will be:

$$A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} \tag{2}$$

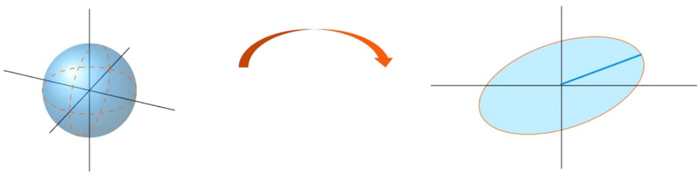

Figure 2: How the unit sphere changes to an ellipse on a 2D plane under the linear transformation of matrix A

Simply finding the eigenvalues of the matrix A is not going to work, as non-square matrices do not have eigenvalues. We can instead turn this problem into a constraint optimization problem. We consider the unit sphere a constraint $||\vec{x}|| = 1$. We then look for the minimum and maximum values a given vector gets after applying the linear transformation A. In other words we want to maximize $||A\vec{x}||$. This is the same as maximizing $||A\vec{x}||^2$. Hence the problem becomes:

$$\text{Maximize } ||A\vec{x}||^2 \text{ under the constraint } \vec{x} = 1.$$

This is just a quadratic form which we know how to solve. The maximum occurs at the maximum value of the eigenvalue of the matrix A^TA, which is square by definition of the matrix product and thus has eigenvalues. Somce we spolved for $||A\vec{x}||^2$ rather then for $||A\vec{x}||$ the semi-major and semi-minor axis become:

$$\sqrt{\lambda_{max}} = 6\sqrt{10} \tag{3}$$
$$\sqrt{\lambda_{min}} = 3\sqrt{10} \tag{4}$$

For every $m \times n$ matrix A all eigenvalues of the symmetric $A^T A$ are non-negative. To prove thsi consider the following: λ is an eigenvalue of $A^T A$ corresponding to $\vec{v}$ with the cosntraint that $||\vec{v}|| = 1$. Then:

$$\lambda = \vec{v}^T \lambda \vec{v} = \vec{v}^T A^T A \vec{v} = A\vec{v} \cdot A\vec{v} = ||A\vec{v}||^2 \geq 0 \tag{5}$$

These square roots of the eigenvalues of $A^T A$ are called the singular values of A. They are defiend as follows: The singular values of an $m \times n$ matrix A are the square roots of the eigenvalues of $A^T A$, usually denoted as $\sigma_1, \sigma_2, \cdots, \sigma_n$ and arranged in decreasing order.

12.2 Orthogonal basis of Col(A)

Suppose that $\lambda_1 \geq \cdots \geq \lambda_r > 0$ and that $\lambda_{r+1} = \cdots = \lambda_n = 0$ are the eigenvalues of$A^T A$. Let $\{\vec{v}_1, \cdots, \vec{v}_n\}$ be an orthonormal set corresponding to the eigenvalues. Then $\{A\vec{v}_1, \cdots, A\vec{v}_n\}$ is an orthogonal basis for the basis of Col(A). Hence r, the number of non-zero singular values of A, is equal to the rank(A) = dim(Col(A)).

12.3 The singular value decomposition

Let A be an $m \times n$ matrix of rank r. Then $A = U\Sigma V^T$, where:

- U is an orthogonal $m \times m$ matrix
- V is an orthogonal $n \times n$ matrix
- Σ is an $m \times n$ matrix consisting of the non-zero singular values of A and 0 entries.

$$\Sigma = \left[\begin{array}{ccc|c} \sigma_1 & \cdots & 0 & \\ \vdots & \ddots & \vdots & O \\ 0 & \cdots & \sigma_r & \\ \hline & O & & O \end{array}\right] \tag{6}$$

To find V we find an orthogonal diagonalization of $A^T A$, arranging its eigenvalues in decreasing order. If $A^T A = PDP^T$ we define $P = V$. For the matrix U:

$$\vec{u}_i = \frac{1}{\sigma_i} A\vec{v}_i, \quad \text{for } i = 1, \cdots, r \tag{7}$$

$\vec{v}_i$ refers to the i-th column of V. We extend $\{\vec{u}_1 \cdots vecu_r\}$ to be an orthonormal basis $\{\vec{u}_1 \cdots vecu_m\}$ for $\mathbb{R}^m$ and define the columns of U as:

$$U = \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_m \end{bmatrix} \tag{8}$$

if $A = U\Sigma V^T$ is a singular value decomposition of A, then $A^T = V\Sigma U^T$ is a singular value decomposition of A^T. Since A and A^T share the same characteristic polynomial they also share the same eigenvalues and by extension the same singular values. The columns of V form an orthonormal set of eigenvectors of $A^T A$ and the columns of U for an orthonormal set of eigenvectors of AA^T.

12.4 The four fundamental subspaces

if $A = U\Sigma V^T$ is a singular value decomposition of A, then:

- $\{\vec{u}_1, \cdots, \vec{u}_r\}$ is an orthonormal basis for Col(A)
- $\{\vec{u}_{r+1}, \cdots, \vec{u}_m\}$ is an orthonormal basis for Nul(A^T)=(Col$(A))^\perp$
- $\{\vec{v}_1, \cdots, \vec{v}_r\}$ is an orthonormal basis for Col(A^T)
- $\{\vec{u}_{r+1}, \cdots, \vec{v}_n\}$ is an orthonormal basis for Nul(A)