***Class Notes*, Math 554, Autumn 2012**

Lecture XIX: The Singular Value Decomposition

1 Singular Value Decomposition

We have now seen the importance of the eigenvalues in determining the "essence" of a linear transformation by identifying its invariant subspaces, and finding the effect of the restriction of the transformation on these subspaces. But what about rectangular matrices? We need a different kind of measure, one that will offer a slightly different characterization: singular values.

Given $A \in \mathcal{M}_{m,n}\mathbb{C}$ (or $\mathcal{M}_{m,n}(\mathbb{R})$), we have the following important theorem. Assume wlog $m \geq n$; if the reverse is true, everything below can be transposed to yield the desired result.

Theorem 1. *(SVD Decomposition) There exist matrices $U \in \mathcal{M}_m(\mathbb{C})$ (repectively, $\mathcal{M}_m(\mathbb{R})$), U unitary, $V \in \mathcal{M}_n(\mathbb{C})$ (repectively, $\mathcal{M}_n(\mathbb{R})$), V unitary, and Σ a diagonal, real matrix in $\mathcal{M}_{m,n}(\mathbb{R})$ with $\sigma_i := \Sigma_{ii} \geq 0$, for all $i \leq n$, and $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n$, such that $A = U\Sigma V^*$. Moreover, $||A||_a = \sqrt{\rho(A^*A)} = \sigma_1$.*

We will give a direct proof of this theorem, based on *a priori* knowledge of the form of the decomposition we are looking for. There are other proofs among which there is an inductive one which closely ressembles the proof for the Schur form; they are a little less intuitive, and we choose to omit them.

Proof. We begin by noting that, as mentioned before, A^*A is a Hermitian, positive semi-definite matrix, therefore it is diagonalizable via an orthogonal transformation: $A^*A = V\Lambda V^*$ for some unitary $n \times n$ V, and $\Lambda \geq 0$, in the sense that all eigenvalues of A^*A are positive. We can assume that the eigenvalues are ordered in non-increasing order on the diagonal of Λ; write $\Lambda = \Sigma^T\Sigma$ for a diagonal matrix $\Sigma \in \mathcal{M}_{m,n}$.

Let k be such that $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_k > 0$ and $\sigma_{k+1} = \ldots = \sigma_n = 0$, and let u_i for $i = 1, 2,\ ldots, k$ be such that $Av_i = \sigma_i u_i$, where $V = [v_1, v_2, \ldots, v_n]$ (v_i is the ith column of V). Let us now examine

$$u_i^* u_j = \frac{1}{\sigma_i \sigma_j} v_i^* A^* A v_i = \delta_{ij}\ ;$$

this indicates that $\{u_1, \ldots, u_k\}$ are an orthonormal set.

Note that for $\sigma_{k+1}, \ldots, \sigma_n$, $v_{k+1}, \ldots, v_n$ are an orthonormal basis for $\mathcal{N}(A)$–we leave it as a simple exercise to note that if $v^*A^*Av = 0$, then $v \in \mathcal{N}(A)$.

Therefore we choose to complete the basis $\{u_1, \ldots, u_k\}$ to an orthonormal basis for the entire space $\mathbb{C}^m$ by tacking on the vectors $\{u_{k+1}, \ldots, u_m\}$, so that the matrix $U = [u_1, \ldots, u_m]$ is unitary.

Thus, $AV = U\Sigma$, and $A = U\Sigma V^*$. □

Note that we could have done the same thing starting not with A^*A, but with AA^*; the fact that the underlying singular values (at least the non-zero ones) would still have been the same is non-trivial, and so we give the following Proposition to explain it.

Proposition 1. *Let $A \in \mathcal{M}_{m,n}$, $B \in \mathcal{M}_{n,m}$. Then AB and BA have the same eigenvalues, except that the larger matrix has an extra $|m - n|$ zero eigenvalues.*

Proof. You have proved this in the case when $m = n$ in HW1 (more or less). For $m \neq n$, we define the following two $(m + n) \times (m + n)$ matrices:

$$C_1 = \begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix}, \qquad C_2 = \begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix}.$$

We note that the matrix $S = \begin{bmatrix} I & A \\ 0 & I \end{bmatrix}$, with inverse $S^{-1} = \begin{bmatrix} I & -A \\ 0 & I \end{bmatrix}$, provides a similarity transformation between C_1 and C_2, as follows: $C_1 = SC_2S^{-1}$. We leave the details of this as an easy exercise.

The matrix C_1 has as eigenvalues those of AB, and an additional n zeros, while C_2's eigenvalues are those of BA plus an additional n zeros. Since the two sets of eigenvalues must be the same, the conclusion follows. □

Remark 1. *One can apply this proposition to A^*A and AA^* to obtain the result about the singular values of A being the same as for A^*, by alternate means.*

Remark 2. $||A||_2^2 = ||A^*A||$, *for all A. This is due to the fact that the operator 2-norm (induced by the Euclidean norm on vectors) is not changed by multiplication with a unitary matrix. Thus* $||A||_2 = \sigma_1$.

An alternative proof follows the lines of the proof we did for the square case.

The vectors u_i are known as the left singular vectors of A, while the vectors v_j are known as the right singular vectors of A. In general, the SVD is not unique (even if one conditions on the strictly non-increasing ordering of singular values, one could have repeated singular values, zero singular values, having to complete the orthogonal basis, etc.

A better (less choice-y) solution is to seek the "compact" singular value decomposition, $A = U_r\Sigma_rV_r^*$, where each of the three matrices on the right has r columns (where $r = k$ is the rank of A, i.e., the number of non-zero singular values), while U_r has m columns, Σ is square, and V_r has n rows. After conditioning on the non-increasing ordering of the singular values, the non-uniqueness can only come from repeated singular values.