Singular Value Decomposition

Jonathan Monette

November 14, 2009

1 Singular Value Decomposition Theorem

Theorem. *If $A \epsilon \mathbb{R}^{m\times n}$ matrix, then there exists orthogonal matrices:*

$$U = (u_0|\cdots|u_{\text{m-1}})\,\epsilon\,\mathbb{R}^{\text{m}\times\text{m}},\;\; V = (v_0|\cdots|v_{\text{n-1}})\,\epsilon\,\mathbb{R}^{\text{n}\times\text{n}}$$

such that

$$U^{\text{T}}AV = \begin{pmatrix} \sum & 0 \\ 0 & 0 \end{pmatrix},\; where \sum \; is\, a\, p\times p \; diagonol\, matrix\,,\; p = min(m,n)$$

and the elements of $\sum$ are ordered. $\sigma_0 \geq \sigma_1 \geq \cdots \geq \sigma_{p-1} \geq 0$.

Proof. Take a unit length vector $x\,\epsilon\,\mathbb{R}^{\text{n}}$ for which the maximum in the definition of the 2-norm of a matrix is attained.

$$\|A\|_2 = \max \frac{\|Ax\|_2}{\|x\|_2}.$$

Now set $Ax = \sigma y$, where $y\,\epsilon\,\mathbb{R}^{\text{m}}$ and $\sigma = \|A\|_2$ which is a normalization factor that makes y a unit length vector. Because any orthonormal set can be extended to form an orthonormal basis for the whole space, it is possible to find $V_1\,\epsilon\,\mathbb{R}^{\text{n}\times(\text{n-1})}$ and $U_1\,\epsilon\,\mathbb{R}^{\text{m}\times(\text{m-1})}$ such that $V = \begin{pmatrix} x & V_1 \end{pmatrix}\epsilon\,\mathbb{R}^{\text{n}\times\text{n}}$ and $U = \begin{pmatrix} y & U_1 \end{pmatrix}\epsilon\,\mathbb{R}^{\text{m}\times\text{m}}$. If we multiply out $U^{\text{T}}AV$ we get

$$U^{\text{T}}AV = U^{\text{T}}\begin{pmatrix} Ax & AV_1 \end{pmatrix} = U^{\text{T}}\begin{pmatrix} \sigma y & AV_1 \end{pmatrix} = \begin{pmatrix} y^T \\ U_1^T \end{pmatrix}\begin{pmatrix} \sigma y & AV_1 \end{pmatrix} = \begin{pmatrix} \sigma y^T y & y^T Ay \\ \sigma U_1^T y & U_1^T AV_1 \end{pmatrix}$$

Because $y\,\epsilon\,U$ and U is orthonormal, $y^T y = 1$, and by the same consequence $U_1^T y = 0$. So,

$$\begin{pmatrix} \sigma y^T y & y^T Ay \\ \sigma U_1^T y & U_1^T AV_1 \end{pmatrix} = \begin{pmatrix} \sigma & w^T \\ 0 & B \end{pmatrix} = \tilde{A}$$

where $w^T = y^T Ay$ and $B = U_1^T AV_1$.

Now we must show that the row vector $w^T = 0$. We know that the orthogonal transformation $\tilde{A}$ does not change the norm of A, $\left\|\tilde{A}\right\|_2 = \|A\|_2 = \sigma$. Let $\tilde{A}$ operate on the row vector $\begin{pmatrix} \sigma & w^T \end{pmatrix}^T$. Since,

$$\left\|\tilde{A}\begin{pmatrix} \sigma \\ w \end{pmatrix}\right\|_2^2 \geq (\sigma^2 + w^T w)^2$$

We have $\left\|\tilde{A}\right\|_2^2 \geq (\sigma^2 + w^T w)$. But by our assumption $\sigma^2 = \|A\|_2^2 = \left\|\tilde{A}\right\|_2^2$, so we must have $w = 0$ for $(\sigma^2 + w^T w) = \sigma^2$.

We now have proven $\tilde{A} = \begin{pmatrix} \sigma & 0 \\ 0 & B \end{pmatrix}$. A simple induction argument on B will finish off the proof. □

[1]

[1]This proof is from "Matrix Computations" by Gene H. Golub and Charles F. Van Loan, 2nd Ed., but expanded upon to portray a clearer interpretation in my opinion.