Analysis of the SVD

Emily Schwartz

December 2018

Definition 1. ***Eigenspace,*** $E(\lambda, T)$
*Suppose $T \in L(V)$ and $\lambda \in \boldsymbol{F}$. The **eigenspace** of T corresponding to λ, denoted $E(\lambda, T)$, is defined by*

$$E(\lambda, T) = null(T - \lambda I) \tag{1}$$

Remark 1. *Since $\lambda I(0) = \lambda * I(0) = \lambda * 0 = 0$ and also $T(0) = 0$, we must have that the vector 0 is always in $null(T - \lambda I)$, since then $T(0) - \lambda I(0) = 0 - 0 = 0$. If $v \neq 0$, then we have that v is in $null(T - \lambda I)$ if and only if v is an eigenvector of T with eigenvalue λ.*

Remark 2. *If S is any linear map $\in \mathcal{L}(V, W)$, then $null(S)$ is a subspace of V. Thus, $null(T - \lambda I)$ is a vector space. This is the same thing as saying that $E(\lambda, T)$ is a vector space. Thus dim $E(\lambda, T)$ may exist, if $E(\lambda, T)$ is finite-dimensional.*

Definition 2. ***adjoint,*** T^*
*Suppose $T \in \mathcal{L}(V, W)$. The **adjoint** of T is the function $T^* : W \to V$ such that $< Tv, w > = < v, T^* w >$ for every $v \in V$ and $w \in W$.*

Definition 3. ***singular values***
*Suppose $T \in \mathcal{L}(V)$. The **singular values** of T are the eigenvalues of $\sqrt{T^* T}$, with each eigenvalue λ repeated dim $E(\lambda, \sqrt{T^* T})$ times.*

Now we will define the Singular Value Decomposition (SVD) of a linear map, T.

Theorem 1. *Singular Value Decomposition*
Suppose $T \in \mathcal{L}(V)$ has singular values $s_1, ..., s_n$. Then there exist orthonormal bases $e_1, ..., e_n$ and $f_1, ..., f_n$ of V such that

$$Tv = s_1 < v, e_1 > f_1 + ... + s_n < v, e_n > f_n \tag{2}$$

for every $v \in V$.

Now we will also define the Singular Value Decomposition (SVD) of an $n \times p$ matrix X. Any matrix X can be broken down into the form

$$X = U\Sigma W^T, \tag{3}$$

Where Σ is an $n \times p$ matrix whose entries on the main diagonal are all positive numbers, and whose entries off of the main diagonal are all 0. The entries on the main diagonal are called the singular values of X.
U is an $n \times n$ square matrix, the columns of which are orthogonal unit vectors of length n called the left singular values of X.
W is a $p \times p$ square matrix, the columns of which are orthogonal unit vectors of length p called the right singular values of X.

Now we will show a connection between the SVD from Theorem 1 and the SVD of a matrix.
Let $W = (e_1, ..., e_n)$, $U = (f_1, ..., f_n)$, where $e_1, ..., e_n$ and $f_1, ..., f_n$ are both from Theorem 1. Also let

$$\Sigma = \begin{pmatrix} s_1 & & \\ & \ddots & \\ & & s_n \end{pmatrix} \tag{4}$$

where $s_1, ..., s_n$ are the singular values from Theorem 1.

Now also let v be any vector in V, where the V comes again from Theorem 1.
Then we have that $v \ = < v, e_1 > e_1 + ...+ < v, e_n > e_n$.
To see this, note that we have $v = a_1e_1 + ... + a_ne_n$, for some $a_1, ..., a_n \in \mathbf{F}$, since $e_1, ..., e_n$ is a basis of V.
If we take the inner product of both sides of the equation above with e_1, we obtain $< v, e_1 > = < a_1e_1 + ... + a_ne_n, e_1 >$
By the Additivity in the First Slot property of inner products, the RHS is equal to: $< a_1e_1, e_1 > + ... + < a_ne_n, e_1 >$
Then by the Homogeneity in the First Slot property of inner products, this is equal to: $a_1 < e_1, e_1 > + ... + a_n < e_n, e_1 >$
Since $e_1, ..., e_n$ is orthonormal, $< e_j, e_1 >$ is equal to 1 if $e_j = e_1$. On the other hand, if $j \neq 1$, then $< e_j, e_1 >$ is equal to 0. Thus, the RHS is equal to a_1.
So, we have that $< v, e_1 > = \ a_1$
Similarly, we have that $< v, e_2 > = \ a_2$, and so on.
Thus, $v \ = \ a_1e_1 + \ ... \ + a_ne_n \ \ = < v, e_1 > e_1 + \ ... + < v, e_n > e_n$.

So if we right-multiply the matrix W^T by the vector v, we obtain:

$$W^Tv \ = \ \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \big(< v, e_1 > e_1 \ + ... + < v, e_n > e_n\big) = \begin{pmatrix} < v, e_1 > \\ \vdots \\ < v, e_n > \end{pmatrix}$$

But then, if we right-multiply the matrix Σ by $W^T v$, we obtain $\Sigma W^T v \; =$

$$\begin{pmatrix} s_1 & & \\ & \ddots & \\ & & s_n \end{pmatrix} \begin{pmatrix} < v, e_1 > \\ \vdots \\ < v, e_n > \end{pmatrix} \; = \; \begin{pmatrix} s_1 < v, e_1 > \\ \vdots \\ s_n < v, e_n > \end{pmatrix}$$

But then, $U\Sigma W^T v \; =$

$$\begin{pmatrix} f_1, ..., f_n \end{pmatrix} \begin{pmatrix} s_1 < v, e_1 > \\ \vdots \\ s_n < v, e_n > \end{pmatrix} \; = \; s_1 < v, e_1 > f_1 + \; ... \; + s_n < v, e_n > f_n,$$

which matches what the Tv from Theorem 1 equals.

References

[1] Axler (2015): "Linear Algebra Done Right, 3rd Edition".

[2] Math.stackexchange
https://math.stackexchange.com/questions/1023130/singular-value-decomposition-in-axlers-book

[3] Wikipedia