1 Singular Value Decomposition

Let $V, W$ be finite-dimensional inner product spaces and let $\varphi : V \to W$ be a linear transformation. Since the domain and range of $\varphi$ are different, we cannot analyze it in terms of eigenvectors. However, we can use the spectral theorem to analyze the operators $\varphi^* \varphi : V \to V$ and $\varphi \varphi^* : W \to W$ and use their eigenvectors to derive a nice decomposition of $\varphi$. This is known as the singular value decomposition (SVD) of $\varphi$.

**Proposition 1.1** Let $\varphi : V \to W$ be a linear transformation. Then $\varphi^* \varphi : V \to V$ and $\varphi \varphi^* : W \to W$ are self-adjoint positive semidefinite linear operators with the same non-zero eigenvalues.

**Proof:** The self-adjointness and positive semidefiniteness of the operators $\varphi^* \varphi$ and $\varphi \varphi^*$ follows from the exercise characterizing positive semidefinite operators in the previous lecture. Specifically, we can see that for any $w_1, w_2 \in W$, 

$$ \langle w_1, \varphi^* \varphi (w_2) \rangle = \langle w_1, \varphi (\varphi^* (w_2)) \rangle = \langle \varphi^* (w_1), \varphi^* (w_2) \rangle = \langle \varphi \varphi^* (w_1), w_2 \rangle. $$

This gives that $\varphi^* \varphi$ is self-adjoint. Similarly, we get that for any $w \in W$

$$ \langle w, \varphi \varphi^* (w) \rangle = \langle w, \varphi (\varphi^* (w)) \rangle = \langle \varphi^* (w), \varphi^* (w) \rangle \geq 0. $$

This implies that the Rayleigh quotient $R_{\varphi^* \varphi}$ is non-negative for any $w \neq 0$ which implies that $\varphi^* \varphi$ is positive semidefinite. The proof for $\varphi \varphi^*$ is identical (using the fact that $(\varphi^*)^* = \varphi$).

Let $\lambda \neq 0$ be an eigenvalue of $\varphi^* \varphi$. Then there exists $v \neq 0$ such that $\varphi^* \varphi (v) = \lambda \cdot v$. Applying $\varphi$ on both sides, we get $\varphi \varphi^* (\varphi (v)) = \lambda \cdot \varphi (v)$. However, note that if $\lambda \neq 0$ then $\varphi (v)$ cannot be zero (why?) Thus $\varphi (v)$ is an eigenvector of $\varphi^* \varphi$ with the same eigenvalue $\lambda$. 

We can notice the following from the proof of the above proposition.

**Proposition 1.2** Let $v$ be an eigenvector of $\varphi^* \varphi$ with eigenvalue $\lambda \neq 0$. Then $\varphi (v)$ is an eigenvector of $\varphi \varphi^*$ with eigenvalue $\lambda$. Similarly, if $w$ is an eigenvector of $\varphi \varphi^*$ with eigenvalue $\lambda \neq 0$, then $\varphi^* (w)$ is an eigenvector of $\varphi^* \varphi$ with eigenvalue $\lambda$. 

1
We can also conclude the following.

**Proposition 1.3** Let the subspaces $V_\lambda$ and $W_\lambda$ be defined as

$$V_\lambda := \{ v \in V \mid \phi^* \phi(v) = \lambda \cdot v \} \quad \text{and} \quad W_\lambda := \{ w \in W \mid \phi^* \phi(w) = \lambda \cdot w \}.$$ 

Then for any $\lambda \neq 0$, $\dim(V_\lambda) = \dim(W_\lambda)$.

Using the above properties, we can prove the following useful proposition, which extends the concept of eigenvectors to cases when we have $\phi : V \to W$ and it might not be possible to define eigenvectors since $V \neq W$ (also $\phi$ may not be self-adjoint so we may not get orthonormal eigenvectors).

**Proposition 1.4** Let $\sigma_1^2 \geq \sigma_2^2 \geq \cdots \geq \sigma_r^2 > 0$ be the non-zero eigenvalues of $\phi^* \phi$, and let $v_1, \ldots, v_r$ be a corresponding orthonormal eigenvectors (since $\phi^* \phi$ is self-adjoint, these are a subset of some orthonormal eigenbasis). For $w_1, \ldots, w_r$ defined as $w_i = \phi(v_i) / \sigma_i$, we have that

1. $\{w_1, \ldots, w_r\}$ form an orthonormal set.
2. For all $i \in [r]$, $\phi(v_i) = \sigma_i \cdot w_i$ and $\phi^*(w_i) = \sigma_i \cdot v_i$.

**Proof:** For any $i, j \in [r], i \neq j$, we note that

$$\langle w_i, w_j \rangle = \left\langle \frac{\phi(v_i)}{\sigma_i}, \frac{\phi(v_j)}{\sigma_j} \right\rangle = \frac{1}{\sigma_i \sigma_j} \cdot \langle \phi(v_i), \phi(v_j) \rangle = \frac{1}{\sigma_i \sigma_j} \cdot \langle \phi^* \phi(v_i), v_j \rangle = \frac{\sigma_i}{\sigma_j} \cdot \langle v_i, v_j \rangle = 0.$$ 

Thus, the vectors $\{w_1, \ldots, w_r\}$ form an orthonormal set. We also get $\phi(v_i) = \sigma_i \cdot w_i$ from the definition of $w_i$. For proving $\phi^*(w_i) = v_i$, we note that

$$\phi^*(w_i) = \phi^* \left( \frac{\phi(v_i)}{\sigma_i} \right) = \frac{1}{\sigma_i} \cdot \phi^* \phi(v_i) = \frac{\sigma_i^2}{\sigma_i} \cdot v_i = \sigma_i \cdot v_i,$$

which completes the proof. 

The values $\sigma_1, \ldots, \sigma_r$ are known as the (non-zero) singular values of $\phi$. For each $i \in [r]$, the vector $v_i$ is known as the right singular vector and $w_i$ is known as the left singular vector corresponding to the singular value $\sigma_i$.

**Proposition 1.5** Let $r$ be the number of non-zero eigenvalues of $\phi^* \phi$. Then,

$$\text{rank} (\phi) = \dim(\text{im}(\phi)) = r.$$ 

2
Using the above, we can write $\phi$ in a particularly convenient form. We first need the following definition.

**Definition 1.6** Let $V, W$ be inner product spaces and let $v \in V, w \in W$ be any two vectors. The outer product of $w$ with $v$, denoted as $|w\rangle \langle v|$, is a linear transformation from $V$ to $W$ such that

$$|w\rangle \langle v| (u) := \langle v, u \rangle \cdot w.$$

In matrix form, over the reals, the outer product of $w$ with $v$ is the rank-1 matrix $wv^T$, as opposed to the inner product $w^Tv$. And then the statement is that $(wv^T)u = w(v^Tu)$.

Note that if $\|v\| = 1$, then $|w\rangle \langle v| (v) = w$ and $|w\rangle \langle v| (u) = 0$ for all $u \perp v$.

**Exercise 1.7** Show that for any $v \in V$ and $w \in W$, we have

$$\text{rank} (|w\rangle \langle v|) = \dim (\text{im} (|w\rangle \langle v|)) = 1.$$

We can then write $\phi : V \to W$ in terms of outer products of its singular vectors.

**Proposition 1.8** Let $V, W$ be finite dimensional inner product spaces and let $\phi : V \to W$ be a linear transformation with non-zero singular values $\sigma_1, \ldots, \sigma_r$, right singular vectors $v_1, \ldots, v_r$, and left singular vectors $w_1, \ldots, w_r$. Then,

$$\phi = \sum_{i=1}^r \sigma_i \cdot |w_i\rangle \langle v_i|.$$

**Exercise 1.9** If $\phi : V \to V$ is a self-adjoint operator with $\dim(V) = n$, then the real spectral theorem proves the existence of an orthonormal basis of eigenvectors, say $\{v_1, \ldots, v_n\}$ with corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$. Check that in this case, we can write $\phi$ as

$$\phi = \sum_{i=1}^n \lambda_i \cdot |v_i\rangle \langle v_i|.$$

Note that while the above decomposition has possibly negative coefficients (the $\lambda_i$s), the singular value decomposition only has positive coefficients (the $\sigma_i$s).