

## Lecture 8: October 24, 2019

Lecturer: Yury Makarychev

## 1 Singular Value Decomposition

Let  $V, W$  be finite-dimensional inner product spaces and let  $\varphi : V \rightarrow W$  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  $\varphi^* \varphi : V \rightarrow V$  and  $\varphi \varphi^* : W \rightarrow W$  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  $\varphi : V \rightarrow W$  be a linear transformation. Then  $\varphi^* \varphi : V \rightarrow V$  and  $\varphi \varphi^* : W \rightarrow W$  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  $\mathcal{R}_{\varphi \varphi^*}$  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  $\lambda$ .*

We can also conclude the following.

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

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

*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  $\varphi : V \rightarrow W$  and it might not be possible to define eigenvectors since  $V \neq W$  (also  $\varphi$  may not be self-adjoint so we may not get orthonormal eigenvectors).

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

1.  $\{w_1, \dots, w_r\}$  form an orthonormal set.
2. For all  $i \in [r]$

$$\varphi(v_i) = \sigma_i \cdot w_i \text{ and } \varphi^*(w_i) = \sigma_i \cdot v_i.$$

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

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

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

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

which completes the proof. ■

The values  $\sigma_1, \dots, \sigma_r$  are known as the (non-zero) singular values of  $\varphi$ . 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  $\varphi^* \varphi$ . Then,*

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

Using the above, we can write  $\varphi$  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

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

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  $\varphi : V \rightarrow W$  in terms of outer products of its singular vectors.

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

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

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

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

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). Why is this the case?