Lecture 7: October 17, 2023

Lecturer: Madhur Tulsiani

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 φ 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 φ. This is known as the singular value decomposition (SVD) of φ.

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 $\mathcal{R}_{\varphi\varphi^*}$ 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 φ 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 λ. ■

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 λ. 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 λ.*

We will use these properties to develop a simple way of understanding the action of linear transformations $\varphi : V \to W$, mapping one inner product space to another. We can also conclude the following.

Proposition 1.3 *Let the subspaces V_λ and W_λ 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 $\varphi : V \to W$ and it might not be possible to define eigenvectors since $V \neq W$ (also φ 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 $\varphi^*\varphi$, and let $v_1, \ldots, v_r$ be a corresponding orthonormal eigenvectors (since $\varphi^*\varphi$ is self-adjoint, these are a subset of some orthonormal eigenbasis). For $w_1, \ldots, w_r$ defined as $w_i = \varphi(v_i)/\sigma_i$, we have that*

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

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

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

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

Thus, the vectors $\{w_1, \ldots, w_r\}$ 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 $\sigma_1, \ldots, \sigma_r$ are known as the (non-zero) singular values of φ. 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 σ_i.

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

$$\mathrm{rank}(\varphi) \; = \; \dim(\mathrm{im}(\varphi)) \; = \; r\,.$$

Using the above, we can write φ 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*

$$\mathrm{rank}\left(|w\rangle \langle v|\right) \; = \; \dim\left(\mathrm{im}\left(|w\rangle \langle v|\right)\right) \; = \; 1\,.$$

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

Proposition 1.8 *Let V, W be finite dimensional inner product spaces and let $\varphi : V \to W$ 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,*

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

Exercise 1.9 *If $\varphi : V \to V$ 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 φ 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 λ_is), the singular value decomposition only has positive coefficients (the σ_is). Why is this the case?

2 Singular Value Decomposition for matrices

Using the previous discussion, we can write matrices in convenient form. Let $A \in \mathbb{C}^{m \times n}$, which can be thought of as an operator from $\mathbb{C}^n$ to $\mathbb{C}^m$. Let $\sigma_1, \ldots, \sigma_r$ be the non-zero singular values and let $v_1, \ldots, v_r$ and $w_1, \ldots, w_r$ be the right and left singular vectors respectively.

Note that $V = \mathbb{C}^n$ and $W = \mathbb{C}^m$ and $v \in V, w \in W$, we can write the operator $|w\rangle \langle v|$ as the matrix wv^*, there v^* denotes $\overline{v^T}$. This is because for any $u \in V$, $wv^*u = w(v^*u) = \langle v, u\rangle \cdot w$. Thus, we can write

$$A \;=\; \sum_{i=1}^{r} \sigma_i \cdot w_i v_i^* \,.$$

Let $W \in \mathbb{C}^{m\times r}$ be a matrix with $w_1, \ldots, w_r$ as columns, such that i^{th} column equals w_i. Similarly, let $V \in \mathbb{C}^{n\times r}$ be a matrix with $v_1, \ldots, v_r$ as the columns. Let $\Sigma \in \mathbb{C}^{r\times r}$ be a diagonal matrix with $\Sigma_{ii} = \sigma_i$. Then, check that the above expression for A can also be written as

$$A \;=\; W\Sigma V^* \,,$$

where $V^* = \overline{V^T}$ as before.

We can also complete the bases $\{v_1, \ldots, v_r\}$ and $\{w_1, \ldots, w_r\}$ to bases for $\mathbb{C}^n$ and $\mathbb{C}^m$ respectively and write the above in terms of unitary matrices.

Definition 2.1 *A matrix $U \in \mathbb{C}^{n\times n}$ is known as a* unitary matrix *if the columns of U form an orthonormal basis for $\mathbb{C}^n$.*

Proposition 2.2 *Let $U \in \mathbb{C}^{n\times n}$ be a unitary matrix. Then $UU^* = U^*U = \mathsf{id}$, where* id *denotes the identity matrix.*

Let $\{v_1, \ldots, v_n\}$ be a completion of $\{v_1, \ldots, v_r\}$ to an orthonormal basis of $\mathbb{C}^n$, and let $V_n \in \mathbb{C}^{n\times n}$ be a unitary matrix with $\{v_1, \ldots, v_n\}$ as columns. Similarly, let $W_m \in \mathbb{C}^{m\times m}$ be a unitary matrix with a completion of $\{w_1, \ldots, w_r\}$ as columns. Let $\Sigma' \in \mathbb{C}^{m\times n}$ be a matrix with $\Sigma'_{ii} = \sigma_i$ if $i \le r$, and all other entries equal to zero. Then, we can also write

$$A = W_m \Sigma' V_n^* \,.$$