Mathematical Toolkit **Autumn 2021**

Lecture 6: October 14, 2021

Lecturer: Madhur Tulsiani

1 Rayleigh quotients: eigenvalues as optimization

Definition 1.1 *Let $\varphi : V \to V$ be a self-adjoint linear operator and $v \in V \setminus \{0_V\}$. The* Rayleigh quotient *of φ at v is defined as*

$$\mathcal{R}_{\varphi}(v) \;:=\; \frac{\langle v, \varphi(v)\rangle}{\|v\|^2}\,.$$

Proposition 1.2 *Let $\dim(V) = n$ and let $\varphi : V \to V$ be a self-adjoint operator with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Then,*

$$\lambda_1 \;=\; \max_{v \in V \setminus \{0_V\}} \mathcal{R}_{\varphi}(v) \quad \textit{and} \quad \lambda_n \;=\; \min_{v \in V \setminus \{0_V\}} \mathcal{R}_{\varphi}(v)$$

Using the above, Rayleigh quotients can be used to prove the spectral theorem for Hilbert spaces, by showing that the above maximum[1] is attained at a point in the space, and defines an eigenvalue if the operator φ is "compact". A proof can be found in these notes by Paul Garrett [**?**].

Proposition 1.3 (Courant-Fischer theorem) *Let $\dim(V) = n$ and let $\varphi : V \to V$ be a self-adjoint operator with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Then,*

$$\lambda_k \;=\; \max_{\substack{S \subseteq V \\ \dim(S) = k}} \min_{v \in S \setminus \{0_V\}} \mathcal{R}_{\varphi}(v) \;=\; \min_{\substack{S \subseteq V \\ \dim(S) = n-k+1}} \max_{v \in S \setminus \{0_V\}} \mathcal{R}_{\varphi}(v)\,.$$

Definition 1.4 *Let $\varphi : V \to V$ be a self-adjoint operator. Φ is said to be* positive semidefinite *if $\mathcal{R}_{\varphi}(v) \geq 0$ for all $v \neq 0$. Φ is said to be* positive definite *if $\mathcal{R}_{\varphi}(v) > 0$ for all $v \neq 0$.*

[1] Strictly speaking, we should write sup and inf instead of max and min until we can justify that max and min are well defined. The difference is that sup and inf are defined as limits while max and min are defined as actual maximum and minimum values in a space, and these may not always exist while we are at looking infinitely many values. Thus, while $\sup_{x \in (0,1)} x \;=\; 1$, the quantity $\max_{x \in (0,1)} x$ does not exist. However, in the cases we consider, the max and min will always exist (since our spaces are closed under limits) and we will use max and min in the class to simplify things.

Proposition 1.5 *Let $\varphi : V \to V$ be a self-adjoint linear operator. Then the following are equivalent:*

1. $\mathcal{R}_\varphi(v) \geq 0$ *for all* $v \neq 0$.
2. *All eigenvalues of* φ *are non-negative.*
3. *There exists* $\alpha : V \to V$ *such that* $\varphi = \alpha^*\alpha$.

The decomposition of a positive semidefinite operator in the form $\varphi = \alpha^*\alpha$ is known as the Cholesky decomposition of the operator. Note that if we can write φ as $\alpha^*\alpha$ for any $\alpha : V \to W$, then this in fact also shows that φ is self-adjoint and positive semidefinite.

2 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 2.1 *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 2.2 *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 can also conclude the following.

Proposition 2.3 *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 2.4 *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

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

Thus, the vectors $\{w_1, \ldots, w_r\}$ 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 2.5 *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 2.6 *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 2.7 *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 2.8 *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 2.9 *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?