Mathematical Toolkit **Autumn 2023**

Lecture 3: October 3, 2023

Lecturer: Madhur Tulsiani

1 Kernel and Image of a Linear Transformation

Definition 1.1 *Let $\varphi : V \to W$ be a linear transformation. We define its* kernel *and* image *as:*

- $\ker(\varphi) \;:=\; \{v \in V \mid \varphi(v) = 0_W\}$.
- $\mathrm{im}(\varphi) \;=\; \{\varphi(v) \mid v \in V\}$.

Proposition 1.2 $\ker(\varphi)$ *is a subspace of V and* $\mathrm{im}(\varphi)$ *is a subspace of W.*

Proposition 1.3 (rank-nullity theorem) *If V is a finite dimensional vector space and $\varphi : V \to W$ is a linear transformation, then*

$$\dim(\ker(\varphi)) + \dim(\mathrm{im}(\varphi)) \;=\; \dim(V)\,.$$

$\dim(\mathrm{im}(\varphi))$ is called the rank and $\dim(\ker(\varphi))$ is called the nullity of φ.

Example 1.4 *Consider the matrix A which defines a linear transformation from $\mathbb{F}_2^7$ to $\mathbb{F}_2^3$:*

$$A \;=\; \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}.$$

- $\dim(\mathrm{im}(\varphi)) = 3$.
- $\dim(\ker(\varphi)) = 4$.
- *Check that* $\ker(\varphi)$ *is a* code *which can recover from one bit of error.*
- *Check that this is also true for the $(2^k - 1) \times k$ matrix A_k where the i^{th} column is the number i written in binary (with the most significant bit at the top).*

This code is known as the Hamming Code and the matrix A is called the parity-check matrix of the code.

2 Eigenvalues and Eigenvectors

Definition 2.1 *Let V be a vector space over the field $\mathbb{F}$ and let $\varphi : V \to V$ be a linear transformation. $\lambda \in \mathbb{F}$ is said to be an* eigenvalue *of φ if there exists $v \in V \setminus \{0_V\}$ such that $\varphi(v) = \lambda \cdot v$. Such a vector v is called an* eigenvector *corresponding to the eigenvalue λ. The set of eigenvalues of φ is called its* spectrum*:*

$$\mathrm{spec}(\varphi) = \{\lambda \mid \lambda \textit{ is an eigenvalue of } \varphi\} \ .$$

Example 2.2 *Consider the matrix*

$$M = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix},$$

which can be viewed as a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$. Note that

$$\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 \\ 0 \end{bmatrix} = \lambda \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

is only satisfied if $\lambda = 0, x_1 = 0$ or $\lambda = 2, x_2 = 0$. Thus $\mathrm{spec}(M) = \{0, 2\}$.

Example 2.3 *It can also be the case that $\mathrm{spec}(\varphi) = \emptyset$, as witnessed by the rotation matrix*

$$M_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

when viewed as a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$.

Example 2.4 *Consider the following transformations:*

- *Differentiation is a linear transformation on the class of (say) infinitely differentiable real-valued functions over $[0,1]$ (denoted by $C^\infty([0,1], \mathbb{R})$). Each function of the form $c \cdot \exp(\lambda x)$ is an eigenvector with eigenvalue λ. If we denote the transformation by φ_0, then $\mathrm{spec}(\varphi_0) = \mathbb{R}$.*
- *We can also consider the transformation $\varphi_1 : \mathbb{R}[x] \to \mathbb{R}[x]$ defined by differentiation i.e., for any polynomial $P \in \mathbb{R}[x]$, $\varphi_1(P) = dP/dx$. Note that now the only eigenvalue is 0, and thus $\mathrm{spec}(\varphi) = \{0\}$.*
- *Consider the transformation $\varphi_{\text{left}} : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$. Any geometric progression with common ratio r is an eigenvector of φ_{left} with eigenvalue r (and these are the only eigenvectors for this transformation).*

Proposition 2.5 *Let $U_\lambda = \{v \in V \mid \varphi(v) = \lambda \cdot v\}$. Then for each $\lambda \in \mathbb{F}$, U_λ is a subspace of V.*

Note that $U_\lambda = \{0_V\}$ if λ is not an eigenvalue. The dimension of this subspace is called the geometric multiplicity of the eigenvalue λ.

Proposition 2.6 *Let $\lambda_1, \ldots, \lambda_k$ be distinct eigenvalues of φ with associated eigenvectors $v_1, \ldots, v_k$. Then the set $\{v_1, \ldots, v_k\}$ is linearly independent.*

Proof: We can prove via induction that for all $r \in [k]$, the subset $\{v_1, \ldots, v_r\}$ is independent. The base case follws from the fact that $v_1 \neq 0_V$, and thus $\{v_1\}$ is a linearly independent set. For the induction step, assume that that the set $\{v_1, \ldots, v_r\}$ is linearly independent.

If the set $\{v_1, \ldots, v_{r+1}\}$ is linearly *dependent*, there exist scalars $c_1, \ldots, c_{r+1} \in \mathbb{F}$ such that

$$c_1 \cdot v_1 + \cdots + c_{r+1} \cdot v_{r+1} \;=\; 0_V \,.$$

Also, note that we must have at least one of $c_1, \ldots, c_r \neq 0$ (since $v_{r+1} \neq 0$). Applying φ on both sides gives

$$\lambda_1 \cdot c_1 \cdot v_1 + \cdots + \lambda_{r+1} \cdot c_{r+1} \cdot v_{r+1} \;=\; 0_V \,.$$

Multiplying the first equality by λ_{r+1} and substracting the two gives

$$(\lambda_1 - \lambda_{r+1}) \cdot c_1 \cdot v_1 + \cdots (\lambda_r - \lambda_{r+1}) c_r \cdot v_r \;=\; 0_V \,.$$

Since all the eigenvalues are distinct, and at least one of $c_1, \ldots, c_r$ is non-zero, the above shows that $\{v_1, \ldots, v_r\}$ is linearly dependent, which contradicts the inductive hypothesis. Thus, the set $v_1, \ldots, v_{r+1}$ must be linearly independent. ■

Definition 2.7 *A transformation $\varphi : V \to V$ is said to be* diagonalizable *if there exists a basis of V comprising of eigenvectors of φ.*

Example 2.8 *The linear transformation defined by the matrix*

$$M \;=\; \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix},$$

is diagonalizable since there is a basis for $\mathbb{R}^2$ formed by the eigenvectors $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Example 2.9 *Any linear transformation $\varphi : V \to V$, with k distinct eigenvalues, where $k = \dim(V)$, is diagonalizable. This is because the corresponding eigenvectors $v_1, \ldots, v_k$ with distinct eigenvalues will be linearly independent, and since they are k linearly independent vectors in a space with dimension k, they must form a basis.*

Exercise 2.10 *Recall that* $\mathrm{Fib} = \left\{ f \in \mathbb{R}^{\mathbb{N}} \mid f(n) = f(n-1) + f(n-2)\ \forall n \geq 2 \right\}$. *Show that $\varphi_{\mathrm{left}} : \mathrm{Fib} \to \mathrm{Fib}$ is diagonalizable. Express the sequence by $f(0) = 1, f(1) = 1$ and $f(n) = f(n-1) + f(n-2)\ \forall n \geq 2$ (known as Fibonacci numbers) as a linear combination of eigenvectors of φ_{left}.*

3 Inner Products

For the discussion below, we will take the field $\mathbb{F}$ to be $\mathbb{R}$ or $\mathbb{C}$ since the definition of inner products needs the notion of a "magnitude" for a field element (these can be defined more generally for subfields of $\mathbb{R}$ and $\mathbb{C}$ known as Euclidean subfields, but we shall not do so here).

Definition 3.1 *Let V be a vector space over a field $\mathbb{F}$ (which is taken to be $\mathbb{R}$ or $\mathbb{C}$). A function $\mu : V \times V \to \mathbb{F}$ is an inner product if*

- *The function $\mu(u, \cdot) : V \to \mathbb{F}$ is a linear transformation for every $u \in V$.*
- *The function satisfies $\mu(u, v) = \overline{\mu(v, u)}$.*
- *$\mu(v, v) \in \mathbb{R}_{\geq 0}$ for all $v \in V$ and is 0 only for $v = 0_V$.*

We write the inner product corresponding to μ as $\langle u, v \rangle_\mu$.

Strictly speaking, the inner product should always be written as $\langle u, v \rangle_\mu$, but we usually omit the μ when the function is clear from context (or we are referring to an arbitrary inner product).

Remark 3.2 *It follows from the first and second properties above, that while the linear transformation $\mu(u, \cdot) : V \to \mathbb{F}$ is linear, the transformation $\mu(\cdot, v) : V \to \mathbb{F}$ defined by fixing the second input, is "anti-linear" or "conjugate-linear" satisfying*

$$\mu(u_1 + u_2, v) = \mu(u_1, v) + \mu(u_2, v) \quad \text{and} \quad \mu(c \cdot u, v) = \bar{c} \cdot \mu(u, v)\,.$$

Example 3.3 *The following are all examples of inner products:*

- *The function $\int_{-1}^{1} f(x)g(x)dx$ for $f, g \in C([-1, 1], \mathbb{R})$ (space of continuous functions from $[-1, 1]$ to $\mathbb{R}$).*
- *The function $\int_{-1}^{1} \frac{f(x)g(x)}{\sqrt{1-x^2}}dx$ for $f, g \in C([-1, 1], \mathbb{R})$.*
- *For $x, y \in \mathbb{R}^2$, $\langle x, y \rangle = x_1y_1 + x_2y_2$ is the usual inner product. Check that $\langle x, y \rangle = 2x_1y_1 + x_2y_2 + x_1y_2/2 + x_2y_1/2$ also defines an inner product.*