Mathematical Toolkit **Autumn 2015**

Lecture 3: October 5, 2015

Lecturer: Madhur Tulsiani

1 Puzzle

Here is another puzzle problem:

Let $A \in \mathbb{F}_2^{k \times n}$, for $n = 2^k - 1$ be a matrix such that the i^{th} column of A equals the number i written in binary (say with the most significant bit at the top). Show that $\ker(A)$ is a *code* which can recover from one-bit errors i.e., if we transmit $x \in \ker(A) \subseteq \mathbb{F}_2^n$ and one of the bits got changed during transmission, the receiver can still find x.

This 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*:*

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

Example 2.2 *Consider the following transformations:*

- *Differentiation is a linear transformation on the class of infinitely differentiable functions and each function of the form $c \cdot \exp(\lambda x)$ is an eigenvector with eigenvalue λ.*
- *Consider the transformation $\varphi_{left} : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$. 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 2.3 *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 2.4 *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.*

Definition 2.5 *A transformation $\varphi : V \to V$ is said to be* diagonalizable *if there exists a basis of V comprising of eigenvectors of φ.*

Exercise 2.6 *Recall that $Fib = \left\{f \in \mathbb{R}^{\mathbb{N}} \mid f(n) = f(n-1) + f(n-2)\ \forall n \geq 2\right\}$. Show that $\varphi_{left} : Fib \to Fib$ 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 $\mu(\cdot, w) : V \to \mathbb{F}$ is a linear transformation for every $w \in V$.*
- *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 $\langle u, v\rangle$.

Strictly speaking, the inner product should 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).

Example 3.2 *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.*

Exercise 3.3 *Let $C > 4$. Check that*

$$\mu(f, g) \;=\; \sum_{n=0}^{\infty} \frac{f(n) \cdot g(n)}{C^n}$$

defines an inner product on the space Fib.

An inner product also defines a norm $\|v\| = \sqrt{\langle v, v\rangle}$ and a hence a notion of distance between two vectors in a vector space. This can be used to define convergence of sequences, and to define infinite sums and limits of sequences (which was not possible in an abstract vector space). However, it might still happen that the limit of a sequence of vectors in the vector space, which converges according to the norm defined by the inner product, may not converge to a vector in the space. Consider the following example.

Example 3.4 *Consider the vector space $C([-1,1],\mathbb{R})$ with the inner product defined by $\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)dx$. Consider the sequence of functions:*

$$f_n(x) = \begin{cases} -1 & x \in \left[-1, \frac{-1}{n}\right) \\ nx & x \in \left[\frac{-1}{n}, \frac{1}{n}\right) \\ 1 & x \in \left[\frac{1}{n}, 1\right] \end{cases}$$

One can check that $\|f_n - f_m\|^2 = O(\frac{1}{n})$ for $m \geq n$. Thus, the sequence converges. However, the limit point is a discontinuous function not in the inner product space. To fix this problem, one can essentially include the limit points of all the sequences in the space (known as the completion of the space). An inner product space in which all (Cauchy) sequences converge to a point in the space is known as a Hilbert space. Many of the theorems we will prove will generalize to Hilbert spaces though we will only prove some of them for finite dimensional spaces.

Definition 3.5 *Two vectors u, v in an inner product space are said to be* orthogonal *if $\langle u, v\rangle = 0$. A set of vectors $\{w_1, \ldots, w_n\}$ is said to be* orthonormal *if $\langle w_i, w_j\rangle = 0$ for all $i \neq j$ and $\|w_i\| = 1$ for all $i \in [n]$.*

Proposition 3.6 *A set $S \subseteq V$ consisting of mutually orthogonal vectors is linearly independent.*

Proposition 3.7 (Gram-Schmidt orthogonalization) *Given a finite set $\{v_1, \ldots, v_n\}$ of linearly independent vectors, there exists a set of orthonornmal vectors $\{w_1, \ldots, w_n\}$ such that*

$$\mathit{Span}(\{w_1, \ldots, w_n\}) \;=\; \mathit{Span}(\{v_1, \ldots, v_n\}) \;.$$

Proof: By induction. The case with one vector is trivial. Given the statement for k vectors and orthonormal $\{w_1, \ldots, w_k\}$ such that

$$\text{Span}(\{w_1, \ldots, w_k\}) \;=\; \text{Span}(\{v_1, \ldots, v_k\}) \;,$$

define

$$u_{k+1} \;=\; v_{k+1} - \sum_{i=1}^{k} \langle v_{k+1}, w_i\rangle \cdot w_i \quad \text{and} \quad w_{k+1} \;=\; \frac{u_{k+1}}{\|u_{k+1}\|}$$

It is easy to check that the set $\{w_1, \ldots, w_{k+1}\}$ satisfies the required conditions. ∎

Corollary 3.8 *Every finite dimensional inner product space has an orthonormal basis.*

In fact, Hilbert spaces also have orthonormal bases (which are countable). The existence of a maximal orthonormal set of vectors can be proved by using Zorn's lemma, similar to the proof of existence of a Hamel basis for a vector space. However, we still need to prove that a maximal orthonormal set is a basis. This follows because we define the basis slightly differently for a Hilbert space: instead of allowing only finite linear combinations, we allow infinite ones. The correct way of saying this is that is we still think of the span as the set of all *finite* linear combinations, then we only need that for any $v \in V$, we can get arbitrarily close to v using elements in the span (a converging sequence of finite sums can get arbitrarily close to it's limit). Thus, we only need that the span is *dense* in the Hilbert space V. However, if the maximal orthonormal set is not dense, then it is possible to show that it cannot be maximal. Such a basis is known as a Hilbert basis.