Lecture : Linear Codes (Contd.) and Examples Thereof

In this lecture, we shall continue with our study of structured, linear codes that are subspaces of $\mathbb{F}^n$, for some finite field $\mathbb{F} = \mathbb{F}_q$ and blocklength n.

We first revisit two simple codes we had seen much earlier:

(i) Repetition code : Let $\mathbb{F} = \mathbb{F}_2$. The repetition code $\mathcal{C}^{REP}$ consists of just two codewords: $\{\underline{0}, \underline{1}\} \subseteq \mathbb{F}_2^n$.

Note that $\mathcal{C}^{REP}$ is an $[n, 1, n]_2$ linear code over $\mathbb{F}_2$.

(ii) Single parity-check code : Let $\mathbb{F} = \mathbb{F}_2$. The single parity-check code $\mathcal{C}^{SPC}$ is the collection:

$$\mathcal{C}^{SPC} = \{\underline{c} : \sum_{i=1}^{n} c_i = 0\}.$$

Claim 1: $\mathcal{C}^{SPC}$ is a linear code with $\dim(\mathcal{C}^{SPC}) = n-1$.

Proof: Linearity is easy to check: indeed, pick $\underline{c}_1, \underline{c}_2 \in \mathcal{C}^{SPC}$ and note that $\underline{c}_1 + \underline{c}_2 \in \mathcal{C}^{SPC}$. By a fact about subspaces we had seen earlier, it follows that $\mathcal{C}^{SPC}$ is linear.

Note also that there are exactly 2^{n-1} codewords in $\mathcal{C}^{SPC}$, since there are as many sequences of "even parity", i.e., the # of 1s in

which is even. Hence, it follows that $\dim(\mathcal{L}^{spc}) = n-1$. ▨

HW: What is $d_{min}(\mathcal{L}^{spc})$?

Def 1: A "generator matrix" G for an $[n,k]_q$ code $\mathcal{L}$ is a $k \times n$ matrix over $\mathbb{F}_q$ whose rows form a basis for $\mathcal{L}$, i.e.,

$$G = \begin{bmatrix} \leftarrow g_1 \rightarrow \\ \leftarrow g_2 \rightarrow \\ \vdots \\ \leftarrow g_k \rightarrow \end{bmatrix}$$

where $\mathcal{L} = \text{span}(g_1, g_2, \cdots, g_k)$.

Remarks: (i) Note that $\text{rank}_{\mathbb{F}}(G) = \dim_{\mathbb{F}}(\mathcal{L})$.

(ii) Since there are likely to be several bases for $\mathcal{L}$, there will be several generator matrices for $\mathcal{L}$.

We use generator matrices for encoding a message to a codeword in the following way:

An encoder for $\mathcal{L}$ is a 1-1 map from message vectors $\underline{m} \in \mathbb{F}_q^k$ to codewords $\underline{c} \in \mathcal{L}$. Since every codeword $\underline{c} \in \mathcal{L}$ can be uniquely expressed as $\underline{c} = \sum_{i=1}^{k} \alpha_i \cdot g_i$, given G, we define the encoder to be the mapping

$$\underline{m} = (m_1, \cdots, m_k) \xmapsto{\mathcal{E}_G} \underline{m}G = \sum_{i=1}^{k} m_i g_i .$$

Remark: Note that since G_1 is full rank, the mapping above is 1-1. Specifically, $E_{G_1}(\underline{0}) = \underline{0}$, for all generator matrices G_1.

Def 2: A generator matrix G_1 for a linear code $\mathcal{L}$ is "systematic," if it is of the form

$$G_1 = [I_k \mid B],$$

where I_k is the $k \times k$ identity matrix.

Remark: (i) Via encoding using a systematic generator matrix, we obtain that

$$\underline{m} \xrightarrow{E_{G_1}} [\underline{m} \mid \underline{m}B]$$

↓ first k symbols of the codeword constitute the message itself. These are called "information symbols."

(ii) Not every code has a systematic generator matrix. Indeed, consider the code

$$\mathcal{L} = \left\{ \begin{matrix} 0000 \\ 0001 \\ 0100 \\ 0101 \end{matrix} \right\}$$

Note that $\mathcal{L}$ is a $[4, 2, 1]$ linear code. The first symbol is 0 in all codewords; hence $\nexists$ a systematic generator matrix for $\mathcal{L}$.

Def 3: A "parity-check matrix" H for an $[n,k]_q$ linear code $\mathcal{L}$ is an $(n-k)\times n$ matrix over $\mathbb{F}_q$ such that $\mathcal{L} = \text{nullspace}(H)$, i.e.,

$$\mathcal{L} = \{\underline{c} : H\underline{c}^T = \underline{0}\}.$$

Theorem 1: Given a generator matrix G and a parity-check matrix H for a linear code $\mathcal{L}$ over $\mathbb{F}_q$, we have that

$$H \cdot G^T = [\underline{0}],$$

where $\underline{0}$ is the all-zeros matrix

i.e., $H g_i^T = \underline{0}$, $\forall i \in [k]$.

Proof: Since $\mathcal{L} = \text{nullspace}(H)$, and since $g_i \in \mathcal{L}$, $\forall i \in [k]$, we have that $H g_i^T = \underline{0}$, $\forall i \in [k]$. ☒

FACT: Via the rank-nullity thm. and the linear-algebraic fact that a basis of a finite-dimensional vector space is a spanning set of the smallest size, we obtain that the rows of H are linearly independent, i.e., that H is full rank.

HW: Suppose that $\mathbb{F} = \mathbb{F}_2$ and that $G = [I_k | B]$ for a linear code $\mathcal{L}$. Obtain a parity-check matrix for $\mathcal{L}$.

Example: The $[7,4,3]_2$ binary Hamming code. Consider the code $\mathcal{C}$ with parity check matrix

$$H = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

← coll^n of all non-zero vectors in $\mathbb{F}_2^3$ as columns, but permuted to make the code systematic. →

HW: Verify that $\dim(\mathcal{C}) = 4$ and that $d_{\min}(\mathcal{C}) = 3$.