Homework 1: SP&LA Study Group

Guille

November 2023

A general note

None of the solutions to these exercises should be more than a few lines long. If you find yourself writing more than half a page or so for a solution, there's probably a simpler way! (Of course, *finding* the simpler way might not, itself, be so simple.) Problem parts beginning with an **(H)** are harder problems.

1 Matrices

In this problem, we'll discuss some properties of matrices. For this problem, let's fix an $m \times n$ matrix with elements in the field $\mathbf{F}$, written $A \in \mathbf{F}^{m \times n}$. Denote the columns of A by the m-vectors $a_1, \ldots, a_n \in \mathbf{F}^m$. As a reminder, if we have an n-vector $x \in \mathbf{F}^n$, then the matrix-vector product between A and x, which we write as Ax, results in the vector

$$Ax = x_1 a_1 + x_2 a_2 + \cdots + x_n a_n, \tag{1}$$

which is a linear combination of the columns of A with the scalars equaling the entries of the vector x.

Part 1. Show that the matrix-vector product, as defined above, is *linear*, in other words that, for any $\alpha \in \mathbf{F}$ and any $x \in \mathbf{F}^n$,

$$A(\alpha x) = \alpha A x,$$

and for any other vector $y \in \mathbf{F}^n$,

$$A(x + y) = Ax + Ay.$$

(We will make use of this a lot in the proofs of the paper!)

(H) Part 2. Let $f : \mathbf{F}^n \to \mathbf{F}^m$ be any linear function that maps n-vectors to m-vectors; *i.e.*, the function f is *linear* because it satisfies

$$f(\alpha x) = \alpha f(x)$$

and

$$f(x + y) = f(x) + f(y),$$

for any scalar $\alpha \in \mathbf{F}$ and any vectors $x, y \in \mathbf{F}^n$. (Compare this with the definition above!) Show that there exists some matrix $A \in \mathbf{F}^{m \times n}$ such that

$$f(x) = Ax.$$

(The fact that every matrix corresponds to a linear function and every linear function corresponds to a matrix is the reason this field is called *linear algebra*.)

Hint. Note that every vector $x \in \mathbf{F}^n$ can be written as a linear combination of the basis vectors,

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Part 3. Say we have a second linear function $h : \mathbf{F}^m \to \mathbf{F}^k$. Let's define the new function

$$w(x) = h(f(x)).$$

This new function w corresponds to taking the output $f(x)$, which is an m-vector, and passing it through h. Show that w is also linear.

(H) Part 4. From part 2, we know that $w(x) = Dx$ for some matrix $D \in \mathbf{F}^{k \times n}$ as it, too, is a linear function. If $f(x) = Ax$ and $h(y) = By$ for some matrix $B \in \mathbf{F}^{k \times m}$, then what does the matrix D correspond to in terms of A, B, or their corresponding columns?

(If you've seen linear algebra before, this question is 'easy' with outside tools, but you should not use other knowledge of linear algebra for this question! Only the matrix-vector product from the definition in (1) is necessary.)

2 Vector spaces

In this problem, we'll explore some basic vector spaces that we talked about in the lecture.

As a reminder, we say that a set of n-vectors, $V \subseteq \mathbf{F}^n$ is a *vector space* if, for any two vectors in this set, $x, y \in V$, and any two scalars $\alpha, \beta \in \mathbf{F}$, the linear combination of these two vectors (with scalars α, β) are also in the set V; *i.e.*,

$$\alpha x + \beta y \in V.$$

Part 1. From the lecture, the range $\mathcal{R}(A)$ is defined as the set containing all possible linear combinations of the columns of the matrix A. Written in set builder notation, this is

$$\mathcal{R}(A) = \{Ax \mid x \in \mathbf{F}^n\}.$$

Show that the range $\mathcal{R}(A)$ of any matrix $A \in \mathbf{F}^{m \times n}$ is a vector space. (You are welcome to use the solution to problem 1, part 1, even if you haven't solved it.)

Part 2. Using the above, the paper shows in §1.1 that the set of evaluations of polynomials of degree at most s on a fixed set of points is itself a vector space. In particular, if we fix a set of points $\alpha_1, \ldots, \alpha_m \in \mathbf{F}$ and we define the set of vectors

$$V = \{(f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_m)) \in \mathbf{F}^m \mid f \text{ is a polynomial of degree} \leq n-1\}, \qquad (2)$$

then this set V is a vector space. Flesh out the proof in the paper to show that this is indeed a vector space! (As a side note, we will make use of this property in our proof of the security of FRI.)

(H) Part 3. The paper also discusses the fact that, for every vector space V, there exists a *parity check matrix* $C \in \mathbf{F}^{k \times m}$ such that, for some vector $x \in \mathbf{F}^m$, we have that $x \in V$ if, and only if, $Cx = 0$. Write out the parity check matrix corresponding to the vector space defined in (2). Two hints: first, take a look at the definition of the Vandermonde matrix in the paper. Second, look up *Lagrange interpolation*.

(H) Other fun. For bonus points, show that, indeed, every vector space has a parity check matrix. You may assume that every vector space V has a matrix A such that $\mathcal{R}(A) = V$ and that this matrix has linearly independent columns. (See below for a reminder of the definition of linear independence.)

3 The ℓ_0 'norm'

From before, remember that we defined the 'norm' $\|y\|_0$ of a vector y in a finite field as the number of nonzero entries of the vector y. In math, that is

$$\|y\|_0 = |\{i \mid y_i \neq 0\}|.$$

We will show three properties we will use throughout the paper.

Part 1. Show that $\|y\|_0 = 0$ if, and only if, $y = 0$. (This is called *definiteness*.)

Part 2. Show that, for any two vectors $x, y \in \mathbf{F}^n$ the *triangle inequality* holds; *i.e.*,

$$\|x + y\|_0 \leq \|x\|_0 + \|y\|_0.$$

Part 3. Show that the ℓ_0 'norm' is zero-homogeneous; *i.e.*,

$$\|\alpha x\|_0 = \|x\|_0,$$

for any $\alpha \in \mathbf{F}$ that is nonzero, $\alpha \neq 0$. (This should be contrasted with a usual norm over the reals or the complex numbers!)

Part 4. A usual thing to do in linear algebra over the real numbers is to take an *inner product* of two vectors x and y with the same number of elements, n. This is defined as

$$x^T y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

The notion of *orthogonality* (that two nonzero vectors are at a 90-degree angle to each other) plays a big role. Unfortunately this notion of orthogonality is not so clear in finite fields.

Show that there is a finite field $\mathbf{F}$ and a nonzero vector $x \in \mathbf{F}^n$ that is orthogonal to itself. Extend this example to show that, for any finite field $\mathbf{F}$, there is always a nonzero vector $x \in \mathbf{F}^n$, that is orthogonal to itself. (As a hint, note that adding 1 to itself $|\mathbf{F}|$ times results in 0.)

4 Linear independence and distance

In this problem, we will relate the distance of a code with some basic properties of linear independence.

From before, we define the *distance* d of a matrix $G \in \mathbf{F}^{m \times n}$ as

$$d = \min_{x \neq 0} \|Gx\|_0$$

where $\|y\|_0$ denotes the number of nonzero entries of the vector y. Similarly, we say that a matrix G has *linearly independent* columns if its nullspace contains only the zero vector; *i.e.*, if

$$\mathcal{N}(G) = \{0\},$$

where the nullspace $\mathcal{N}(G)$ is defined

$$\mathcal{N}(G) = \{y \in \mathbf{F}^n \mid Gy = 0\}.$$

A basic fact from linear algebra (which you are free to use here) is that any matrix G with linearly independent columns has at least as many rows as it has columns; *i.e.*, since G is an $m \times n$ matrix, then $m \geq n$.

Part 1. Show that a generator matrix G is linearly independent if, and only if, its distance d is positive, $d > 0$.

Part 2. Given a matrix G which has some distance $d > 0$, show that we can remove any $d-1$ rows to get a new matrix $\tilde{G} \in \mathbf{F}^{(m-d+1)\times n}$ that also has linearly independent columns. Argue that this means that the distance must satisfy

$$d \leq m - n + 1.$$

(This bound on the distance is known as the *Singleton bound* in coding theory.) Codes that achieve this bound at equality are known as 'maximum-distance separable' or MDS codes.

As a fun sidenote: if G is a generator matrix for the Reed–Solomon code (see §1.3.2), then we know that $d = m - n + 1$, making the Reed–Solomon code an MDS code. In a sense, it is the highest-distance code for the chosen dimensions (message length and block size).

5 (Probabilistic?) Implications

In this section, we'll explore both 'traditional' logic implications and probabilistic logic implications. As a reminder, given two statements P and Q, we say P *implies* Q when

$$\neg(P \wedge \neg Q). \tag{3}$$

As a second reminder, given random variables r and r' which are taken from some (known) distribution and statements P_r and $Q_{r'}$, each depending on the randomness of r and r', then we say that

$$P_r \underset{p}{\implies} Q_{r'}$$

whenever $\Pr(P_r \wedge \neg Q_{r'}) \leq p$.

Part 1. Show that if P implies Q and Q implies T, then P implies T using the definition of implication given above in (3). You will need to assume the *law of excluded middle*: either Q or $\neg Q$. This is called the *transitivity* of implication.

(It is actually possible to have logic without the law of excluded middle, but implications must be defined differently for transitivity to hold. Lucky for us, we only deal with finite—if very large—sets so we can forget any of this conversation ever happened.)

Part 2. Prove that the probabilistic implications have a similar transitivity property; in particular, if $P_r \underset{p}{\implies} Q_{r'}$ and $Q_{r'} \underset{p'}{\implies} T_{r''}$ then $P_r \underset{p+p'}{\implies} Q_{r'}$. (This proof is provided in appendix A of the paper, but you are encouraged to only peek if you need a hint :)

Part 3. A common special case will be to take Q to be a deterministic event (*i.e.*, the statement does not depend on any randomness). This is true in many ZK protocols as we want to ensure some logical (*i.e.*, deterministic) statement is true but somehow 'reduce' this claim down to an easier-to-check claim over a smaller statement with some randomness.

First, show that in this case, $P_r \underset{p}{\implies} Q$ is the same as the statement

$$\Pr(P_r) \leq p$$

when $\neg Q$. (In some way, we may think of this as: the probability of P_r, given $\neg Q$, is less than p. If p is very small, like 2^{-80}, yet we observe P_r, then it is very unlikely that $\neg Q$.)

Second, let's say we have n statements, all depending on the same randomness r, given by $P_r^1 \underset{p}{\implies} Q^1$, $P_r^2 \underset{p}{\implies} Q^2$, until, $P_r^n \underset{p}{\implies} Q^n$. Show that

$$P_r^1 \wedge P_r^2 \wedge \cdots \wedge P_r^n \underset{p}{\implies} Q^1 \wedge Q^2 \wedge \cdots \wedge Q^n.$$

(We use a special case of this fact in the proof in §3.1.2 in the paper.) *Hint.* This should be *very* simple. If you find yourself writing more than a few lines, then it's likely you're overthinking it!