6.440 Essential Coding Theory Feb 20, 2008

Lecture 5

Lecturer: Madhu Sudan *Scribe: MinJi Kim*

1 Lecture overview

In this lecture, we will take a detour from coding theory for a crash course in algebra, which will be useful in the future. As the *crash course* part suggests, this lecture will not provide an extensive coverage of algebra, but only the little part of it that we really need.

- Definitions
- Polynomial rings
- Finite fields

2 Definitions

A *ring* is a set $\mathcal{R}$ with two binary operations usually called *addition*, denoted as $+$, and *multiplication*, denoted as $\cdot$, such that $(\mathcal{R}, +)$ satisfies the five axioms of closure, associativity, commutativity, identity element (called *zero*, 0), and inverse element; and $(\mathcal{R}, \cdot)$ satisfies the three axioms of closure, associativity, and identity element (called *one*, 1). Furthermore, multiplication $(\cdot)$ distributes over addition $(+)$, i.e. $a \cdot (b + c) = a \cdot b + a \cdot c$.

Rings are commutative over addition $(a + b = b + a)$, but need not be commutative over multiplication $(a \cdot b = b \cdot a)$. Rings that satisfy commutativity for multiplication are called *commutative rings*. In this lecture, we will only consider commutative rings.

A *field* $\mathbb{F}$ is a ring where every non-zero element has a multiplicative inverse.

3 Polynomial rings

A *polynomial ring* is a set of polynomials in one or more variables with coefficients from a ring. As the name suggests, the polynomial ring with addition $(+)$ and multiplication $(\cdot)$ itself forms a ring.

To be more precise, let $\mathcal{R}$ be a ring. A polynomial $P(X)$ is defined to be of the form:

$$P(X) = a_0 + a_1 X + ... + a_{n-1} X^{n-1} + a_n X^n$$

where $a_0, a_1, ..., a_n \in \mathcal{R}$. The *degree* of a polynomial, $deg(P(X))$, is the index of the highest non-zero coefficient. A polynomial $P(X)$ is *monic* if its highest coefficient is one.

Consider two polynomials $P_1(X) = a_0 + a_1X + ... + a_{n-1}X^{n-1} + a_nX^n$ and $P_2(X) = b_0 + b_1X + ... + b_{n-1}X^{n-1} + b_nX^n$. Then, addition and multiplication of these two polynomials are given by the following formulas:

$$P_1(X) + P_2(X) = \sum_{i=0}^{n}(a_i + b_i)X^i$$

and

$$P_1(X) \cdot P_2(X) = \sum_{i=0}^{2n}\left(\sum_{j+k=i} a_jb_k\right)X^i.$$

It is not hard to check that this set of polynomials with coefficients from a ring $\mathcal{R}$ with operations $+$ and $\cdot$ itself forms a ring, which is denoted by $R[X]$. In this lecture, we will only consider polynomial rings where the coefficients are from a field $\mathbb{F}$, and we will denote this polynomial ring as $\mathbb{F}[X]$.

There are many useful properties of polynomials, of which we discuss three of them here.

- We can consider a polynomial $P(X) = a_0 + a_1X + ... + a_{n-1}X^{n-1} + a_nX^n$ as a finite sequence of elements from $\mathbb{F}$, since $P(X)$ is defined by the coefficients $a_0, a_1, ..., a_d \in \mathcal{R}$.
- Given a polynomial $P(X) = \sum_i a_iX^i \in \mathbb{F}[X]$ and an element $\alpha \in \mathbb{F}$, we define $P(\alpha) = \sum_i a_i\alpha^i$. This mapping $\mathbb{F}[X] \times \mathbb{F} \to \mathbb{F}$ is called the *evaluation map.*
- An element $\alpha \in \mathbb{F}$ is a *root* of $P(X)$ if $P(\alpha) = 0$.
- Given polynomials $P_1(X) = \sum_{i=0}^{n} a_iX^i$ and $P_2(X) = \sum_{j=0}^{m} b_jX^j$ where $n \geq m$, $P_1(X)$ is equivalent to $P_2(X)$ if $a_i = b_i$ for all $i \leq m$ and $a_i = 0$ for all $i > m$. For example, the two polynomials $1 + X + 0X^2$ is equivalent to $1 + X$, which agrees with our intuition.

Lastly, we present the division algorithm. Although it is called an "algorithm", this is actually a theorem that states the outcome of the process of division of polynomials.

Theorem 1 (Division Algorithm) *Given any two $f, g \in \mathbb{F}[X]$, there exists unique $q, r \in \mathbb{F}[X]$ such that $deg(r) < deg(g)$ and $f = q \cdot g + r$.*

The proof of the theorem consists of two parts: existence proof and uniqueness proof. The existence of q and r can be proven by long division, and the uniqueness can be shown by contradiction.

3.1 Unique factorization domain (UFD)

A unique factorization domain is a commutative ring $\mathcal{R}$ in which every element is either *reducible*, if it can be written as a product of other elements, or *irreducible*. This representation is unique in the sense that if $x \in \mathcal{R}$ can be presented as $x = p_1 p_2 ... p_n$ and $x = q_1 q_2 ... q_m$ where p_i's and q_j's are irreducible elements of $\mathcal{R}$, then $m = n$ and there is a permutation $\Pi : [1, n] \rightarrow [1, n]$ such that $p_i = q_{\Pi(i)}$.

Any field $\mathbb{F}$ is trivially a UFD, since every non-zero element has a multiplicative inverse. In addition, if $\mathcal{R}$ is a UFD, then so is $\mathcal{R}[X]$. Therefore, $\mathbb{F}[X]$ is a UFD.

3.2 Fundamental theorem of algebra

Lemma 2 (Fundamental Theorem of Algebra) *A non-zero degree d polynomial $P(X) \in \mathbb{F}[X]$ has at most d roots.*

The proof of this lemma follows from the Division Algorithm. To give an idea of how the proof works, lets assume that α_1 is a root. Then, $(X - \alpha_1)$ divides $P(X)$. This implies that $P(X) = Q(X) \cdot (X - \alpha_1) + 0$. By induction, if there are d roots $\alpha_1, \alpha_2, ..., \alpha_d$, then $P(X) = \overline{Q}(X) \cdot (X - \alpha_1) \cdot (X - \alpha_2) \cdot ... \cdot (X - \alpha_d)$ where $deg(\overline{Q}(X)) \leq 0$. This bounds the number of roots of $P(X)$ to at most d.

The corollary of the Fundamental Theorem is very useful, and we state it below.

Corollary 3 *Let $P(X)$ and $Q(X)$ be distinct degree d polynomials in $\mathbb{F}[X]$. Then, there are at most d points $\alpha \in \mathbb{F}$ such that $P(\alpha) = Q(\alpha)$.*

3.3 Polynomial interpolation

Given some set of data points, polynomial interpolation finds a polynomial that goes through these points. Formally, given a set of data points $(\alpha_i, \beta_i) \in \mathbb{F} \times \mathbb{F}$, $i = 0, 1, ..., d$, interpolation aims to find $P(X) = \sum_{j=0}^{d} c_j X^j$ such that $P(\alpha_i) = \beta_i$ for $i = 0, 1, ..., d$ (or vice versa).

One approach to polynomial interpolation is to use matrices. By putting together the $d + 1$ linear equations into a matrix form, we have:

$$\begin{bmatrix} 1 & \alpha_0 & ... & \alpha_0^d \\ \vdots & \vdots & & \vdots \\ 1 & \alpha_j & ... & \alpha_j^d \\ \vdots & \vdots & & \vdots \\ 1 & \alpha_d & ... & \alpha_d^d \end{bmatrix} \cdot \begin{bmatrix} c_0 \\ \vdots \\ c_j \\ \vdots \\ c_d \end{bmatrix} = \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_j \\ \vdots \\ \beta_d \end{bmatrix}$$

The matrix one the left is a *Vandermonde matrix*, and its determinant is given by $\prod_{1 \leq i < j \leq n} (\alpha_j - \alpha_i)$. By solving this system of equations for c_j, we can construct the polynomial $P(X)$. However, this may cost $O(d^3)$ time to solve.

Another approach to polynomial interpolation problem is to use Lagrange. By choosing the Lagrange basis, we get the identity matrix (instead of the Vandermonde matrix) and this allows us to reduce the cost to $O(d^2)$.

3.4 Multi-variate polynomials

We define a new set $\mathbb{F}(X)$ of ratio of pairs of polynomials as follows:

$$\mathbb{F}(X) = \left\{ \frac{f}{g} \;\middle|\; f, g \in \mathbb{F}[X], g \neq 0 \right\}.$$

We say that $f/g, f'/g' \in \mathbb{F}(X)$ are equivalent if $f' = f \cdot p$ and $g' = g \cdot p$.

It is known that $\mathbb{F}(X)$ is a field, and therefore, $\mathbb{F}(X)[Y]$ is a polynomial ring with variable Y and coefficients from $\mathbb{F}(X)$. As a result, everything we have discussed so far in this lecture applies to $\mathbb{F}(X)[Y]$. Using this new polynomial field, $\mathbb{F}(X)[Y]$, we can now define multi-variate polynomial rings.

For example, to construct a polynomial ring with two variables X and Y, we first construct the polynomial ring $\mathbb{F}(X)$, and then, on top of it, the ring $\mathbb{F}(X)[Y]$. For example, $P(X,Y) = X^2Y^2 + 3XY^2 + 5X^2Y + 4XY - 5X$ is a polynomial in Y with coefficients $\mathbb{F}(X)$ as follows: $(X^2 + 3X)Y^2 + (5X^2 + 4X)Y + (-5X)$.

This can be extended to include n variables $X_1, X_2, ..., X_n$ such that we get n-variate polynomial rings $\mathbb{F}[X_1, X_2, ..., X_n]$. A polynomial $f \in \mathbb{F}[X_1, X_2, ..., X_n]$ is defined to be of the form:

$$f = \sum c_{\overline{d}} X_1^{d_1} X_2^{d_2} ... X_n^{d_n}.$$

The degree of polynomial f is defined to be:

$$deg(f) = \max_{\overline{d}, c_{\overline{d}} \neq 0} \left\{ \sum_{i=0}^{n} d_i \right\}.$$

4 Finite fields

A *finite field* is a field that contains finitely many elements. Let $\mathbb{F}_q$ denote a finite field of q elements. There is a unique field of order $q = p^t$ for every prime p and positive integer t (up to isomorphism). Finite fields will become very handy in this course, and it would be useful to have an efficient representation of $\mathbb{F}_q$.

The most naive way of representing $\mathbb{F}_q$ is by creating two tables of size $q \times q$: a multiplication table and an addition table. However, this requires polynomial in q bits to represent.

4.1 Computable representation

For a more efficient representation, we use the uniqueness of fields of order $q = p^t$ (up to isomorphism). This implies that for a prime field $\mathbb{F}_p = \mathbb{Z}_p$. Therefore,

any multiplication/addition operation on this field can be translated to that of on integers modulo p. This allows us to compact our representation of $\mathbb{F}_p$ to $\log p$ bits.

To extend this to prime-power fields $\mathbb{F}_{p^t}$, we pick an irreducible monic polynomial $H(X)$ of degree exactly t. Then, we define $\mathbb{F}_{p^t} = \{g \in \mathbb{F}_p[X], deg(g) < t\}$ with addition and multiplication in $\mathbb{F}_p[X]$ modulo $H(X)$. This representation of $\mathbb{F}_{p^t}$ needs t elements of $\mathbb{F}_p$, therefore, requires $t \log p = \log q$ bits.

However, to use this representation, we need to find the irreducible monic polynomial $H(X)$ given p and t. This can be done probabilistically in time polynomial in $\log q$ as well as deterministically in time polynomial in (p, t).

4.2 Vector representation

There is a nice correspondence between the finite field $\mathbb{F}_{p^t}$ and vector space $(\mathbb{F}_p)^t$, which respects addition. More formally, there is a correspondence between $\alpha, \beta \in \mathbb{F}_{p^t}$ and $v_\alpha, v_\beta \in (\mathbb{F}_p)^t$ such that $\alpha + \beta$ corresponds to $v_\alpha + v_\beta = v_{\alpha+\beta}$. Unfortunately, there is no clear way to incorporate multiplication into this correspondence; therefore, this vector representation is incomplete.

4.3 Matrix representation

Although vector representation by itself is incomplete, when combined with the matrix representation (which we introduce here) can be quite useful.

Continuing with the notation used above, let $v_\alpha \in (\mathbb{F}_p)^t$ correspond to $\alpha \in \mathbb{F}_{p^t}$. Now, consider a linear map $L_\beta : (\mathbb{F}_p)^t \to (\mathbb{F}_p)^t$ such that $L_\beta(v_\alpha) = v_{\alpha\beta}$. Since L_β is linear, we can represent it using a $t \times t$ matrix $M_\beta \in \mathbb{F}_p^{t \times t}$ such that $L_\beta(v_\alpha) = M_\beta v_\alpha$. From this representation, we have that $M_{\alpha+\beta} = M_\alpha + M_\beta$ and $M_{\alpha\beta} = M_\alpha \cdot M_\beta$.

Note that the space of matrices in $\mathbb{F}_p^{t \times t}$ is much larger than that of our finite field $\mathbb{F}_{p^t}$, since $\mathbb{F}_{p^t}$ has p^t elements where as $\mathbb{F}_p^{t \times t}$ has p^{t^2} elements. Therefore, we only need a subset of these matrices to represent a finite field.