Foundations of Probabilistic Proofs (Fall 2022)
Note 4: Doubly-Efficient IPs
Date: 2022.9.28

This note contains definitions, theorems, facts, *etc.* that are not fully explained in lectures due to limited time. If you think there are anything missing or any mistakes, please contact ziyi.guan@epfl.ch.

In the lecture, we mention low degree extension of polynomials when we present the GKR protocol. This is a concept that will show up repeatedly throughout the semester. We give it a formal treatment in this note.

1 Lagrange polynomials

Given a set of distinct pairs $\{(x_i, y_i)\}_{i\in[n]}$, Lagrange polynomials give a way to compute the interpolation of these points. By interpolation, we mean a polynomial that would evaluate to y_i on x_i for all i. A straightforward way to do interpolation is to find n polynomials, where each of them takes care of one coordinate, and then add them up together. This approach gives us the Lagrange polynomials.

Definition 1 (Lagrange interpolating polynomial)**.** *Given a set of distinct pairs $\{(x_i, y_i)\}_{i\in[n]}$, the* **Lagrange basis** *for polynomials of degree at most n for the pairs is the set of polynomials $\{L_0(x), L_1(x), \ldots, L_n(x)\}$ where*

$$L_i(x) = \prod_{0\le j\le n, j\neq i} \frac{x - x_j}{x_i - x_j}.$$

The **Lagrange interpolating polynomial** *for these pairs is a linear combination of the above polynomials:*

$$L(x) = \sum_{i=0}^{n} y_i L_i(x).$$

We make the following interesting observations for Lagrange polynomials:

- Every Lagrange basis $L_i(x)$ has degree less than $n+1$ (and so is the Lagrange interpolating polynomial). Moreover, $L_i(x_i) = 1$ and $L_i(x_j) = 0$ for all $j \neq i$.
- The Lagrange interpolating polynomial is the unique polynomial of lowest degree that interpolates the set of distinct pairs (x_i, y_i) for $i \in [n]$, because if two polynomials of degree less than $n+1$ agree on $n+1$ points then they are equal by fundamental theorem of algebra.
- $L(x)$ can be evaluated at any point in $p(n)$ operations where p is a polynomial.

Exercise 1. *Let $\mathbb{F}^{<=n}[X]$ be the vector space of polynomials of x of degree $\le n$ with coefficients in the field $\mathbb{F}$, show that the Lagrange polynomials $\{L_i\}_{i\in[n]}$, with $\{a_i\}_{i\in[n]}$ being (fixed) distinct elements in $\mathbb{F}$, form a basis of $\mathbb{F}^{<=n}[X]$, where:*

$$L_i(x) := \prod_{0\le j\le n, j\neq i} \frac{x - a_j}{a_i - a_j}$$

Furthermore, show that coordinates of polynomial $f \in \mathbb{F}^{<=n}[X]$ in this basis are $\{f(a_0), f(a_1), \ldots, f(a_n)\}$.

Solution. We only need to show that $\{L_i\}_{i\in[n]}$ are linearly independent. Suppose that $\sum_{i=0}^{n} b_i L_i(x) = 0$, then setting $x = a_i$ suffices. ■

Example 1. *Consider three points* $(1,1),(2,4),(3,9)$*, then the interpolating polynomial is:*

$$L(x) = 1 \cdot \frac{x-2}{1-2} \cdot \frac{x-3}{1-3} + 4 \cdot \frac{x-1}{2-1} \cdot \frac{x-3}{2-3} + 9 \cdot \frac{x-1}{3-1} \cdot \frac{x-2}{3-2} = x^2$$

2 Low-degree extensions

Low-degree extension is an important technique used in the GKR protocol and will appear in many other protocols throughout the semester. We formally introduce it in this section. Since we are using finite fields throughout the semester, here we give a formal definition of $\mathbb{F}_q$.

Definition 2 (Finite field)**.** *A finite field or Galois field* $\mathbb{F}_q$ *is a field that contains a finite number of elements. This means that multiplication, addition, subtraction and division (excluding division by zero) are defined and satisfy the rules of arithmetic known as the field axioms. The order, denoted by* q*, of a finite field* F_q *is its number of elements, which is either a prime number or a prime power. That is to say,* $q = p^k$ *for some prime number* p *and* $k \in \mathbb{N}^+$*.*

Definition 3 (Low-degree extension)**.** *Suppose* $\mathbb{H}$ *and* $\mathbb{F}$ *are finite fields such that* $\mathbb{F}_2 \subseteq \mathbb{H} \subseteq \mathbb{F}$*. For every function* $f\colon \mathbb{H}^n \to \mathbb{F}$*, a* **low degree extension** *of* f *to* $\mathbb{F}$ *is a (unique)* n*-variable polynomial* $\widehat{f}$ *over* $\mathbb{F}$ *such that the individual degree is less than* $|\mathbb{H}|$ *and* $\widehat{f}|_{\mathbb{H}^n} \equiv f$*.*

One can think of a low degree extension $\widehat{f}$ as an encoding of f in the following sense: if two Boolean functions f and f' disagree at just a single point over $\mathbb{H}^n$, then their low degree extension must differ almost everywhere (given $|\mathbb{H}| \ll |\mathbb{F}|$). This is because, letting $p \coloneqq f - f'$, $\Pr[p(x_1, x_2, \ldots, x_n) = 0] \leq \frac{d}{|\mathbb{F}^n|}$, where d is the total degree of p, by polynomial identity lemma. Hence this property gives the verifier surprising power to check the proof sent by the prover.

We introduce an approach to constructing a low degree extension for a given function and give a careful analysis to show that it conforms with **??**.

- The univariate case ($n = 1$): A low degree extension $\widehat{f}$ of f can be constructed by the Lagrange interpolation of $\{(a, f(a))\}_{a\in\mathbb{H}}$,

  $$\widehat{f}(x) = \sum_{a\in\mathbb{H}} f(a) L_{\mathbb{H},a}(x)$$

  where $L_{\mathbb{H},a}(x) \coloneqq \prod_{b\in\mathbb{H}\setminus\{a\}} \frac{x-b}{a-b}$ is the corresponding Lagrange polynomial over $\mathbb{H}$.

- The multivariate case($n \geq 1$): A generalization to the univariate case, we have

  $$\widehat{f}(x_1, \ldots, x_n) = \sum_{a_1,\ldots,a_n\in\mathbb{H}} f(a_1, \ldots, a_n) \prod_{i=1}^{n} L_{\mathbb{H},a_i}(x_i)$$

  where the product can be rewritten as $L_{\mathbb{H}^n,a_1,\ldots,a_n}(x_1, \ldots, x_n) \coloneqq \prod_{i=1}^{n} L_{\mathbb{H},a_i}(x_i)$.

- $L_{\mathbb{H}^n,a_1,\ldots,a_n}(x_1,\ldots,x_n)$ can be viewed as Lagrange polynomial extened to the multivariate case since $L_{\mathbb{H}^n,a_1,\ldots,a_n}(a_1,\ldots,a_n) = 1$ and $L_{\mathbb{H}^n,a_1,\ldots,a_n}(x_1,\ldots,x_n) = 0$ for all $x \in \mathbb{H}^n$ with $(x_1,\ldots,x_n) \neq (a_1,\ldots,a_n)$. Remark that $L_{\mathbb{H}^n,a_1,\ldots,a_n}$ can be evaluated in time $\mathrm{poly}(|\mathbb{H}|, n)$ and $\widehat{f}$ in time $\mathrm{poly}(|\mathbb{H}|^n)$
- For the special case of equality polynomials, that is, the function $I\colon \mathbb{H}\times\mathbb{H}\to\mathbb{F}$ with $I(x,y) = 1$ if $x = y$ and 0 otherwise. we denoted its low-degree extension by $\widehat{I}_{\mathbb{H}} \in \mathbb{F}(x,y)$. To be explicit,
$$\widehat{I}_{\mathbb{H}}(x,y) = \sum_{a\in\mathbb{H}} L_{\mathbb{H},a}(x)L_{\mathbb{H},a}(y).$$
For any number of variables $n \in \mathbb{N}$, we can extend it to
$$\widehat{I}_{\mathbb{H}^n}(x_1,\ldots,x_n,y_1,\ldots,y_n) = \prod_{i=1}^{n} \widehat{I}_{\mathbb{H}}(x_i,y_i).$$
Notice that the low degree extensions for I seems different from that constructed by Lagrange polynomials, yet they are indeed consistent (see the exercise below). Meanwhile, the polynomial $\widehat{I}_{\mathbb{H}}$ can be evaluated in time $\mathrm{poly}(|\mathbb{H}|)$ and the polynomial $\widehat{I}_{\mathbb{H}^n}$ in time $\mathrm{poly}(|\mathbb{H}|, n)$, which is much faster than low degree extension $\widehat{f}$ for general f.

Exercise 2. *Show that the definition of $\widehat{I}_{\mathbb{H}^n}$ is consistent with the low degree extensions constructed by Lagrange polynomials.*

Example 2. *Consider a function $f\colon \{0,1\}^2 \to \mathbb{F}_5$ with*
$$f(0,0) = 1, f(0,1) = 2, f(1,0) = 1, f(1,1) = 4,$$
then its low degree extension is
$$\begin{aligned}\widehat{f}(x_1,x_2) &= (f(0,0)\cdot\frac{x_1-1}{0-1}\cdot\frac{x_2-1}{0-1} + f(0,1)\cdot\frac{x_1-1}{0-1}\cdot\frac{x_2-0}{1-0} + \\ &\quad f(1,0)\cdot\frac{x_1-0}{1-0}\cdot\frac{x_2-1}{0-1} + f(1,1)\cdot\frac{x_1-0}{1-0}\cdot\frac{x_2-0}{1-0}) \mod 5 \\ &= (2x_1x_2 + x_2 + 1) \mod 5\end{aligned}$$
such that $\widehat{f}(x_1,x_2)$ is over $\mathbb{F}_5$.