Zippel's Sparse Interpolation

November 23, 2023 7:07 PM

1979.

Let $f \in k[x_1, \ldots, x_n]$ e.g. $k = \mathbb{Z}_p$, p large. $p = 2^{31}-1$.

Assume we know $d_i = \deg(f, x_i)$.

If $n = 1$ use ordinary interpolation (Newton or Lagrange) which need $d_1 + 1$ points.

Suppose $n = 3$ and

$$f = 3x_1^2 x_2 + 7x_1^2 x_3^2 + 2x_2^2 x_3 + 7x_2. \quad d_1 = \deg(f, x_1) = 2.$$

① Let S be a large finite subset of k.
If $k = \mathbb{Z}_p$ then $S = \mathbb{Z}_p$.

Pick $\beta_0 \in S$ at random. Say $\beta_0 = 1$.

Interpolate $f(\beta_0 = 1, x_2, x_3) = 3x_2 + 7x_3^2 + 2x_2^2 x_3 + 7x_2$
$= 7x_3^2 + 2x_2^2 x_3 + 10x_2$. recursively.

Since $\deg(f, x_1) = 2$ we need to pick $\beta_1, \beta_2 \in k$ s.t. $\beta_1 \neq \beta_2 \neq \beta_0$.
Compute $f(\beta_1, x_2, x_3)$ and $f(\beta_2, x_2, x_3)$ then interpolate x_1 in f using dense interpolation.

How can we do this efficiently?
Zippel's sparse assumption is

$$f(\beta_i, x_2, x_3) = a_1 \cdot x_3^2 + a_2 x_2^2 x_3 + a_3 \cdot x_2$$

for $a_1, a_2, a_3 \in k$, i.e., we did not lose any monomials in x_2 and x_3 using $x_1 = \beta_0$.

Writing $f = x_3^2(7x_1^2) + x_2^2 x_3(2) + x_2(3x_1^2 + 7)$.
So $\beta_0 = 0$, and $3\beta_0^2 + 7 = 0$ cause missing terms.

Let $f = \sum_{i=1}^{s} a_i(x_1) \cdot M_i(x_2, x_3)$ where $\underset{5}{s} \le \underset{4}{t}$.

Let $h(x_1) = \prod_{i=1}^{s} a_i(x_1) \in k[x_1]$.

Zippel assumes $a_i(\beta_0) \neq 0$. for $1 \le i \le s$.

$$\Pr[\text{a missing term occurs}] = \Pr[h(\beta_0)=0] \le \frac{\deg(h)}{|S|} \le \frac{s\cdot d_1}{|S|} \le \frac{t d_1}{|S|}.$$

(2) $$f = (3x_2^1 + 7x_3^2)x_1^2 + (2x_2^2x_3 + 7x_2).$$

$$f(2,1,3) = (3+63)\cdot 4 + (6+7) = 264+13 = 277.$$

Assumed form $f = a_1 x_3^2 + a_2 x_2^2 x_3 + a_3 x_2$.

Pick $\beta_1 = 2$. We need 3 values of $f(\beta_1, x_2, x_3)$ to determine a_1, a_2, a_3.

$f(x_1, x_2, x_3)$

$x_1=2,\ x_2=1,\ x_3=3$: $\quad 277 = a_1\cdot 9 + a_2\cdot 3 + a_3\cdot 1$

$x_1=2,\ x_2=2,\ x_3=1$: $\quad 74 = a_1 + 4a_2 + 2a_3$

$x_1=2,\ x_2=3,\ x_3=0$: $\quad 57 = 3a_3 \Rightarrow a_3 = 19$

$\Big\}\ a_2 = 2,\ a_1 = 28.$

$$\Rightarrow f(\beta_1=2, x_2, x_3) = 28x_3^2 + 2x_2^2x_3 + 19x_2$$

We needed 3 points instead of $(2+1)(2+1)=9$ points.

Pick $\beta_2 = 3$. Using 3 points again we get

$$\rightarrow f(\beta_2=3, x_2, x_3) = 63x_3^2 + 2x_2^2x_3 + 34x_2$$

recursively $\rightarrow f(\beta_0=1, x_2, x_3) = 7x_3^2 + 2x_2^2x_3 + 10x_2$

$$\rightarrow f(\beta_1=2, x_2, x_3) = 28x_3^2 + 2x_2^2x_3 + 19x_2.$$

Interpolating x_1: $\quad 7x_1^2 \qquad 2 \qquad 3x_1^2+7$

$$\Rightarrow f(x_1,x_2,x_3) = 7x_1^2x_3^2 + 2x_2^2x_3 + 3x_1^2x_2 + 7x_2.$$

We needed $\underbrace{\overset{\beta_1,\beta_2}{2}}_{=d_1}\cdot \overset{s}{3}$ + recursive call instead of $(2+1)(2+1)(2+1) = 27$.

$$\le \underset{x_1}{d_1\cdot t} + \underset{x_2}{d_2\cdot t} + \underset{x_3}{d_3+1} \quad \leftarrow \text{dense.}$$

$$\le t\sum_{i=1}^{n} d_i + 1 \qquad \in O(t\textstyle\sum d_i).$$

If missing terms occur in Zippel's algorithm, it will output some $g \neq f$. How can we check if $g \neq f$ if we have a black-box $B: k^n \to k$ for f?

Pick $\alpha \in \mathbb{Z}_p^n$ at random.

If $g \neq f$ then probably $B(\alpha) = f(\alpha) \neq g(\alpha)$.

But $B(\alpha) = g(\alpha)$ is possible. E.g.

$$f = 2x_1x_2 + \underline{(x_1 - \alpha_1)}x_2^2$$

$$g = 2x_1x_2$$

Let $h = f - g$.

Suppose $f \neq g$.

$Pr[B(\alpha) = g(\alpha)] = Pr[h(\alpha) = 0] \leq \frac{\deg(h)}{p}$ by Schwartz-Zippel