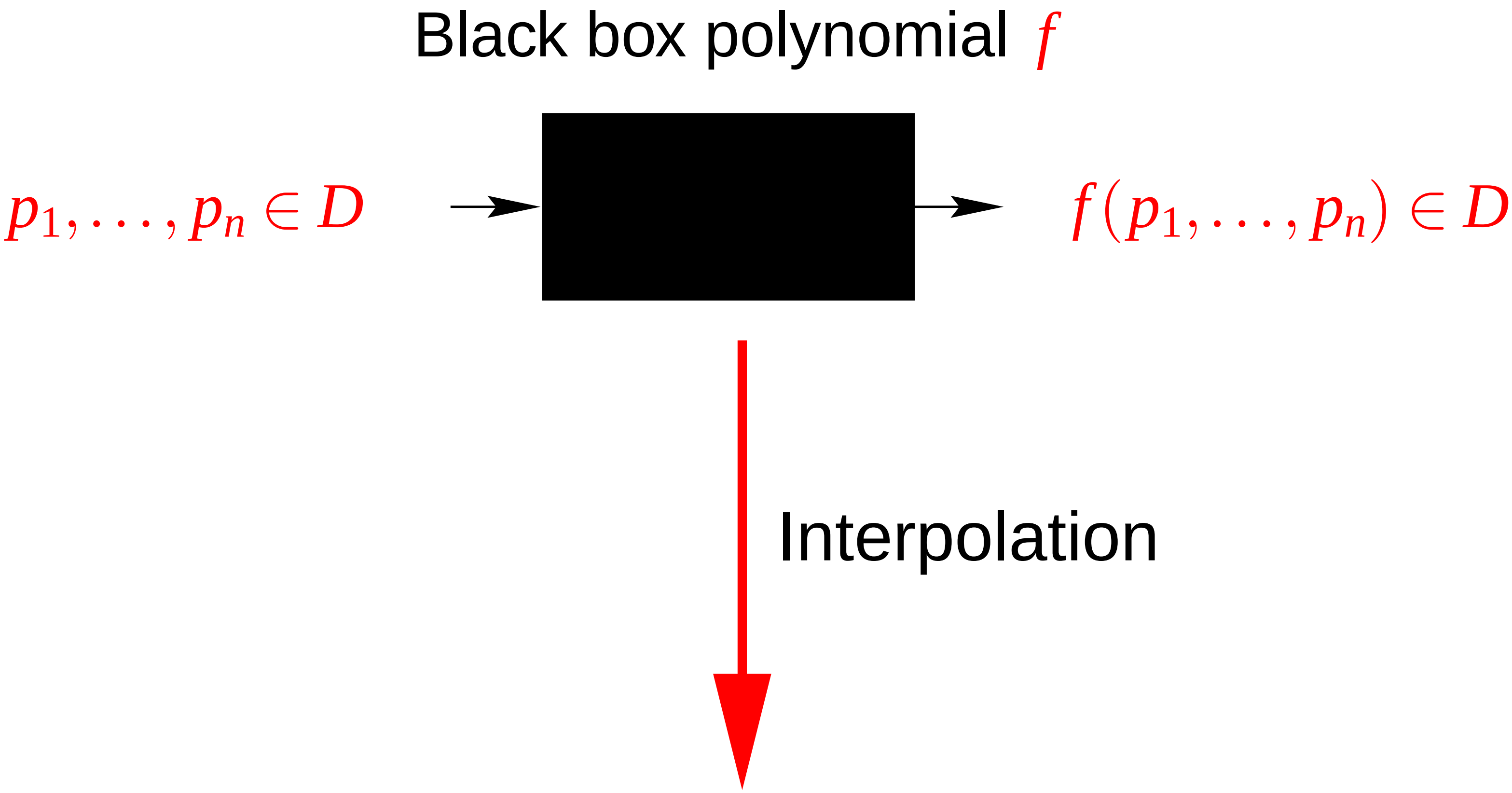
Early Termination in Polynomial Interpolation Algorithms and Applications

Wen-shin Lee
University of Waterloo
www.wen-shin.com

Black box polynomial interpolation



$$f(x_1, \ldots, x_n) = \sum_{i=1}^{t} c_i x_1^{e_{1,i}} \cdots x_n^{e_{n,i}} \in D[x_1, \ldots, x_n]$$

What if $f(x_1, \ldots, x_n)$ is sparse?

When $f(x_1,\ldots,x_n) = \sum_{i=1}^{t} c_i x_1^{e_{1,i}} \cdots x_n^{e_{n,i}}$ is sparse:

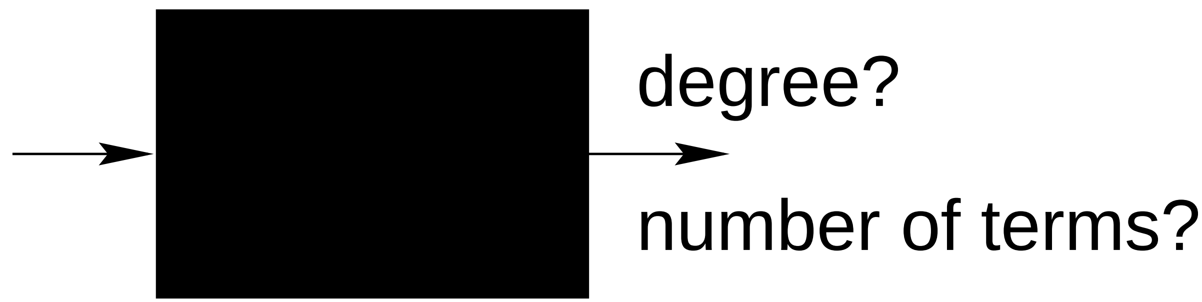
Zippel's probabilistic interpolation (1979)

Need: $d_j \geq \deg f(x_j)$ for $1 \leq j \leq n$

Ben-Or/Tiwari deterministic algorithm (1988)

Need: $T \geq t$

Without T and d_j



- Guess and check

  And double the guess if fails.

- Early termination strategy

  Interpolate the polynomial at a random point, when the polynomial stops changing, it is done with high probability.

Early termination in Newton univariate interpolation

- Interpolate $f(x)$ on random $p_0,\ p_1,\ p_2,\ldots$

$$f^{[i]}(x) = c_0 + c_1(x-p_0) + \cdots + c_i(x-p_0)\cdots(x-p_{i-1})$$

- When $c_i = 0$, $f = f^{[i]}$ and $i = \deg f + 2$ with high probability.

Early termination Ben-Or/Tiwari sparse interpolation

Kaltofen, Lee, Lobo (2000)

Interpolate: $f = \sum_{i=1}^{t} c_i x_1^{e_{1,i}} \cdots x_n^{e_{n,i}}$

- With distinct random p_j, compute minimal linear generator Λ of $f(p_1, \ldots, p_n),\ f(p_1^2, \ldots, p_n^2), \ldots, f(p_1^i, \ldots, p_n^i), \ldots$.

> Berlekamp/Massey algorithm: compute "discrepancy" Δ_i.
> When $\Delta_i = 0$ at $i > 2L$, $i = 2t+1$ and Λ is determined with high probability.

- Recover terms in f by finding roots of Λ.
- Locate coefficients c_i in f.

In $f(x_1, \ldots, x_n) = \sum_{i=1}^{t} c_i x_1^{e_{1,i}} \cdots x_n^{e_{n,i}}$, powers of x_j are revealed as powers of p_j as f evaluated at

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} p_1^i \\ p_2^i \\ \vdots \\ p_n^i \end{bmatrix}, \text{ for } i \geq 1.$$

Polynomial evaluations in any given power basis

When $f(x_1,\ldots,x_n) = \sum_{i=1}^{t} \gamma_i y_1^{\delta_{1,i}} \cdots y_n^{\delta_{n,i}}$, where A, S are given

$$\underbrace{\begin{bmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,n} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \ldots & a_{n,n} \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \underbrace{\begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}}_{S} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

The powers of y_j can be revealed in:

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,n} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \ldots & a_{n,n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix} = \begin{bmatrix} p_1^i \\ p_2^i \\ \vdots \\ p_n^i \end{bmatrix}, \quad \text{for } i \geq 1.$$

Sparse interpolation in any given power basis

Interpolate: $f = \sum_{i=1}^{t} \gamma_i y_1^{\delta_{1,i}} \cdots y_n^{\delta_{n,i}}$

- A power basis Y is given as A and S:

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}_{Y} = \underbrace{\begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \underbrace{\begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}}_{S}$$

- Perform Ben-Or/Tiwari algorithm (early termination) to

$f(x_1, \dots, x_n)$ evaluated at $A^{-1} \begin{bmatrix} p_1^i - s_1 \\ \vdots \\ p_n^i - s_n \end{bmatrix}$ for $i \geq 1$.

Example Interpolate $f(x_1, x_2)$ in the power basis of y_1, y_2 :

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}}_{Y} = \underbrace{\begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \underbrace{\begin{bmatrix} 1 \\ -5 \end{bmatrix}}_{S}$$

At powers of y_j, $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A^{-1} \begin{bmatrix} y_1^i - 1 \\ y_2^i - (-5) \end{bmatrix} = \begin{bmatrix} 2y_1^i - y_2^i - 7 \\ -3y_1^i + 2y_2^i + 13 \end{bmatrix}$,

with random p_1, p_2 perform early termination Ben-Or/Tiwari on

$$\begin{gathered} f(2p_1 - p_2 - 7, -3p_1 + 2p_2 + 13), \\ f(2p_1^2 - p_2^2 - 7, -3p_1^2 + 2p_2^2 + 13), \\ \vdots \\ f(2p_1^i - p_2^i - 7, -3p_1^i + 2p_2^i + 13), \\ \vdots \end{gathered}$$

Early termination sparse interpolation in non-standard bases

Kaltofen, Lee (2002)

Pochhammer basis:

$$f(x) = \sum_{i=1}^{t} c_i x^{\overline{e_i}}$$

$$x^{\overline{n}} = x(x+1)\cdots(x+n-1)$$

Chebyshev basis:

$$f(x) = \sum_{i=1}^{t} c_i T_{\delta_i}(x)$$

$$T_0(x) = 1, \quad T_1(x) = x$$
$$n \geq 2: \quad T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x)$$

Consider the polynomial:

$$f(x_1,\ldots,x_n) = \sum_{i=1}^{t} c_i x_1^{e_{1,i}} \cdots x_n^{e_{n,i}}$$

in shifted basis $y_j = x_j + s_j$:

$$\sum_{i=1}^{t(s)} \gamma_i \underbrace{(x_1+s_1)}_{y_1}^{e_{1,i}} \cdots \underbrace{(x_n+s_n)}_{y_n}^{e_{n,i}}$$

$t(s)$ depends on $s = (s_1,\ldots,s_n)$

Questions:

- Find a sparsest shift of f within set S:

  $s = (s_1,\ldots,s_n) \in S$ and $t(s)$ is minimized.

- T-sparse shifts of f within set S:

  $s = (s_1,\ldots,s_n) \in S$ and $t(s) \leq T$.

Example:

$$f(x_1,x_2) = x_1^5x_2^2 - 6x_1^5x_2 + 9x_1^5 + 10x_1^4x_2^2 - 60x_1^4x_2 + 90x_1^4 + 40x_1^3x_2^2 - 240x_1^3x_2 + 360x_1^3 + 80x_1^2x_2^2 - 480x_1^2x_2 + 720x_1^2 + 80x_1x_2^2 - 480x_1x_2 + 720x_1 + 32x_2^2 - 192x_2 + 288$$

$$= (\underbrace{x_1+2}_{y_1})^5(\underbrace{x_2-3}_{y_2})^2$$

$(2,-3)$ is a sparsest shift of $f(x_1,x_2)$

Shifts in the standard power basis: $A = I_n$

$$\underbrace{\begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}}_{A=I_n} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \underbrace{\begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}}_{\text{shift}} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}_{Y}$$

Early termination Ben-Or/Tiwari sparse interpolation

$\sum_{i=1}^{t} c_i x_1^{e_{1,i}} \cdots x_n^{e_{n,i}}$ $\quad$ $\sum_{i=1}^{t(s)} \gamma_i (x_1 + s_1)^{\delta_{1,i}} \cdots (x_n + s_n)^{\delta_{n,i}}$

$\downarrow \Delta_i$ $\quad$ Berlekamp/Massey $\quad$ $\downarrow \Delta_i$

$\Delta_{2t+1} = 0$ $\quad$ symbolic $x_1, \ldots, x_n$ $\quad$ $\Delta_{2t(s)+1} = 0$

Leave shifts s_j as symbols: $s_j \longrightarrow z_j$

Compute sparsest shifts $s = (s_1, \ldots, s_n)$: solve first $\Delta_i(z) = 0$
for symbolic $x_1, \ldots, x_n$

Minimize: $i = 2t(s) + 1$

Compute sparsest shifts in the standard power basis

Giesbrecht, Kaltofen, Lee (2002)

- Perform fraction-free Berlekamp/Massey algorithm on

$$f(y_1 - z_1, \ldots, y_n - z_n), \ldots, f(y_1^i - z_1, \ldots, y_n^i - z_n), \ldots$$

> The fraction-free Berlekamp/Massey algorithm:
> $\Delta_i(z_1, \ldots, z_n, y_1, \ldots, y_n)$ are polynomials in $z_1, \ldots, z_n, y_1, \ldots, y_n$.

- Solve $z_1, \ldots, z_n$ in $\boxed{\Delta_i = 0}$ for all $y_1, \ldots, y_n$, which minimizes i.

When $f = f(x)$: $\quad f(x) \in \mathbb{Q}[x]$.

Sparsest shifts in any power basis given as A and S

Sparsest shifts for a set of polynomials $f_1, \ldots, f_m \in D[x_1, \ldots, x_n]$:

consider $G(x_1, \ldots, x_n, z_0) = f_1 + z_0 f_2 + \cdots + z_0^{m-1} f_{m-1} + z_0^{m-1} f_m$.

Sparse shifts in Chebyshev, Pochhammer bases:

cf. Kaltofen, Lee (2002)

Sparsify linear transform: A unknown

efficiency? structured A?

Gaspard Clair Franois Marie Riche de Prony

Essai expérimental et analytique sur les lois de la dilatabilité et sur celles de la force expansive de la vapeur de l'eau et de la vapeur de l'alkool, à différentes températures.

J. de l' École Polytechnique 1:24–76, 1795.

For a function $f : \mathbb{R} \to \mathbb{R}$, and $t \in \mathbb{Z}_{>0}$, find c_i, a_i such that

$$f(x) = \sum_{i=1}^{t} c_i e^{a_i x}$$

Prony's method vs. Ben-Or/Tiwari algorithm

Giesbrecht, Labahn, Lee (2002)

Prony	Ben-Or/Tiwari
Interpolate: $f(x) = \sum_{i=1}^{t} c_i e^{a_i x}$	Interpolate: $f(x_1,\ldots,x_n) = \sum_{i=1}^{t} c_i x_1^{e_{1,i}} \cdots x_n^{e_{n,i}}$
1. Solve λ_i, $j = 0,\ldots,t-1$: $\sum_{i=0}^{t-1} \lambda_i f(i+j) = -f(t+j)$ 2. e^{a_i} are zeros of $\Lambda = \theta^t + \lambda_{t-1}\theta^{t-1} + \cdots + \lambda_0$	1. Compute† the minimal Λ that generates* $\{f(p_1^i,\ldots,p_n^i)\}_{i=0}^{2t-1}$ 2. $p_1^{e_{1,i}} \cdots p_n^{e_{n,i}}$ are zeros of $\Lambda = \theta^t + \lambda_{t-1}\theta^{t-1} + \cdots + \lambda_0$
3. Determine c_i from e^{a_i} and evaluations of f	3. Determine c_i from $p_1^{e_{1,i}} \cdots p_n^{e_{n,i}}$ and evaluations of f

† Berlekamp/Massey algorithm

$*$ $p_1,\ldots,p_n$ distinct primes

Applications to Prony's method:

interpolate $f(x) = \sum_{i=1}^{t} c_i e^{a_i x}$

Berlekamp/Massey algorithm

Early termination

Multivariate case

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{t} c_i e^{a_{1,i} x_1 + \cdots + a_{n,i} x_n}$$

Linear differential system

e.g. interpolate f from $f(p), f^{(1)}(p), f^{(2)}(p), \ldots$.

Further extensions