Strong Solution to Smale's 17th Problem for Strongly Sparse Systems

Paula Burkhardt

Pomona College and Texas A&M University

July 23, 2014

Smale's 17th Problem

Smale's 17th Problem

Does there exist a deterministic algorithm which approximates a root of a polynomial system and runs in polynomial time on average?

Approximate Roots

Definition – Approximate Root (Smale [1986])

Suppose $f : \mathbb{C}^n \to \mathbb{C}^n$ *is a multivariate polynomial. Let* $z \in \mathbb{C}^n$ *be a point such that*

$$|\zeta - N_f^k(z)| \leq \frac{1}{2^{2^k} - 1}|\zeta - z|$$

where N_f *is the Newton operator,* $z \mapsto z - Df(z)^{-1}f(z)$*, and* ζ *is an actual root of* f*. Then* z *is an approximate root of* f *with associated true root* ζ*.*

Approximate Roots: γ Theory

Definition – γ (Smale [1986])

For $f : \mathbb{C}^n \to \mathbb{C}^n$ analytic in a neighborhood of $z \in \mathbb{C}^n$ let

$$\gamma(f, z) := \sup_{k \geq 2} \left| \frac{f'(z)^{-1} f^{(k)}(z)}{k!} \right|^{\frac{1}{k-1}}$$

γ Theorem (Smale [1986])

Suppose $f : \mathbb{C}^n \to \mathbb{C}^n$ is analytic in a neighborhood of z containing a root ζ of f and that $f'(\zeta)$ is invertible. If

$$|z - \zeta| \leq \frac{3 - \sqrt{7}}{2\gamma(f, \zeta)}$$

then z is an approximate root of f with associated true root ζ.

Approximate Roots: α Theory

Definition – β and α (Smale [1986])

For $f : \mathbb{C}^n \to \mathbb{C}^n$ analytic in a neighborhood of $z \in \mathbb{C}^n$ let

$$\beta(f, z) := |f'(z)^{-1} f(z)|$$

and

$$\alpha(f, z) := \beta(f, z)\gamma(f, z)$$

α Theorem (Smale [1986])

There exists a universal constant α_0 such that if $z \in \mathbb{C}^n$ with $\alpha(f, z) < \alpha_0$ then z is an approximate root of f.
Smale, 1986: $\alpha_0 \geq 0.1370707$.
Wang and Han, 1989: $\alpha_0 \geq 3 - 2\sqrt{2}$.

Examples of γ Theory

Lemma (B.)

For any univariate polynomial $f(x_1) = c_1 x_1^{a_1} + \ldots + c_t x_1^{a_t}$ where $c_1, \ldots, c_t \in \mathbb{C}^*$ and $a_1, \ldots, a_t \in \mathbb{N}$ with $0 < a_1 < \ldots < a_t$ we have that $\gamma(f, z) \leq \left| \frac{a_t - 1}{2z} \right|$ for all $z \in \mathbb{C}$.

Example

Let $f(x_1) = x_1^d - c$. z is an approximate root of f if $|c| > 1$ and

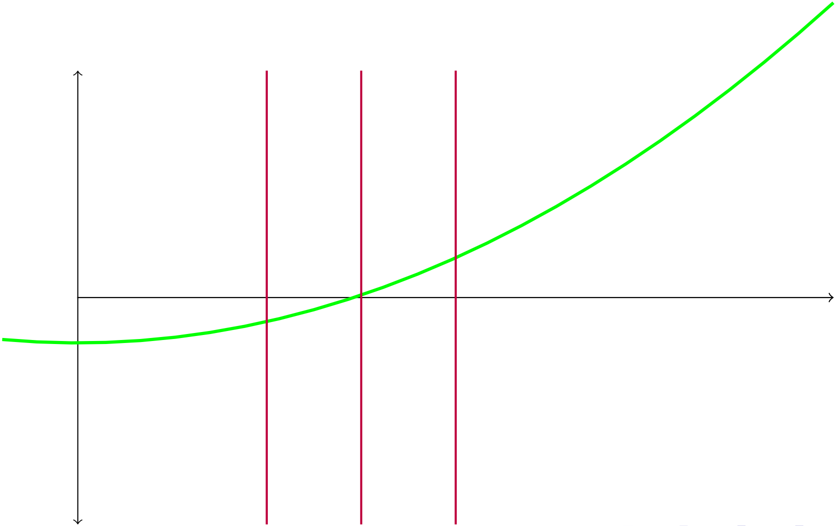
$$|z - c^{\frac{1}{d}}| \leq \frac{1}{3d} \leq \frac{3 - \sqrt{7}}{d - 1} |c^{\frac{1}{d}}|$$

or $0 < c < 1$ and

$$|z - c^{\frac{1}{d}}| \leq \frac{3 - \sqrt{7}}{d} |c| \leq \frac{3 - \sqrt{7}}{d - 1} |c^{\frac{1}{d}}|$$

The Bisection Method

Consider $f(x_1) := x_1^d - c$ where $c > 0$ and $d \in \mathbb{N}$.



The Bisection Method

The complexity of evaluating f at each iteration is $O(\log(d)^2)$ and we need no more than $O(\log(d) \pm \log(c))$ iterations so:

Lemma (B.)

A root of a random binomial of the form $f(x_1) := x_1^d - c$ for $c > 0$ and $d \in \mathbb{N}$ can be approximated in time $O(\log(d)^3)$ on average using the bisection method.

Monic Univariate Binomials

What if c is complex? Let $c = a + bi = re^{i\theta}$ and observe that $c^{\frac{1}{d}} = r^{\frac{1}{d}} e^{\frac{i\theta}{d}}$.

Algorithm for Monic Univariate Binomials

1. *Approximate* $r^{\frac{1}{d}}$ *to within* $\frac{\varepsilon}{5}$ *using bisection. Call this approximation* r_0.
2. *Approximate* θ *by approximating* $\arctan\left(\frac{b}{a}\right)$ *to within* $\frac{d\varepsilon}{5}$ *with Taylor series. Call this approximation* α.
3. *Approximate* $e^{i\frac{\alpha}{d}}$ *to within* $\frac{\varepsilon}{5}$ *via Taylor series. Call the approximations for the cosine and sine components* s_k *and* t_k *respectively.*
4. *Return* $r_0(s_k + it_k)$.

Monic Univariate Binomials

Recall that our approximate root is $r_0(s_k + it_k)$.

- s_k and t_k are kth partial sums where $k = O(\log d)$
- The complexity of computing s_k and t_k is then $O(\log d((\log d)^2 + (\log d)^2(\log\log d)^2))$.

Proposition (B.)

The average complexity of our algorithm is $O((\log d)^3(\log\log d)^2)$: better than polynomial in d.

General Univariate Binomals

Consider $f(x_1) := c_1 x_1^d - c_2$ for $d \in \mathbb{N}$ and $c_1, c_2 \in \mathbb{C}^*$. Note that

$$f(z) = 0 \iff z^d - \frac{c_2}{c_1} = 0$$

so let $c = \frac{c_2}{c_1}$ and apply our algorithm for the monic case.

Binomial Systems

Example

For a diagonal system of binomials $f(x_1, \ldots, x_n) = \begin{cases} x_1^{a_1} - c_1 \\ \vdots \\ x_n^{a_n} - c_n \end{cases}$

and $x = (x_1, \ldots, x_n) \in \mathbb{C}^n$ *we have*

$$\gamma(f, x) \leq \frac{\sqrt{2n} X \max\{|x_i^{-a_i}|\} ||x||_1^{d-2} d^2}{2}$$

where all $a_i \in \mathbb{Z} \setminus \{0\}$, $d = \max\{a_i\}$, $c_i \in \mathbb{C}$, $X = \max\{|x_i|\}$, *and* $||x||_1 = \sqrt{1 + ||x||^2}$.
For a general system of binomials we have

$$\gamma(f, x) \leq \frac{\sqrt{2n^{n+1}} X \max\{|x_i^{-a_i}|\} ||x||_1^{d-2} d^{n+1}}{2}$$

Binomial Systems: Diagonal Systems

Algorithm for Diagonal Binomial Systems

Input: A diagonal binomial system f.

1. *Let ε be an appropriate lower bound on $\frac{3-\sqrt{7}}{2\gamma(f,\zeta)}$ where $\zeta = (\zeta_1, \ldots, \zeta_n)$ is a true root of the system.*
2. *Approximate each ζ_i to within $\frac{\varepsilon}{\sqrt{n}}$ by some α_i.*
3. *Return $\alpha = (\alpha_1, \ldots, \alpha_i)$.*

Lemma (B.)

On average the complexity of this algorithm is $O(n(d \log d)^3 + n(d \log d)^3(\log d + \log\log d))^2)$

Smith Normal Form

Definition –Smith Normal Form

An $n \times n$ nonsingular matrix S is in Smith Normal Form if

1. *It is a diagonal matrix*
2. *All of its entries are positive*
3. *If* $S = \begin{bmatrix} d_1 & 0 & \dots & 0 \\ & \ddots & & 0 \\ 0 & \dots & 0 & d_n \end{bmatrix}$ *then* $d_i \mid d_{i+1} \, \forall i \in \{1, \dots, n\}$.

Example –Smith Normal Form

$$\begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 6 \\ 4 & 8 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

Smith Normal Form

Proposition

For any $n \times n$ *matrix* A *there exists a unique matrix* S *such that* $UAV = S$ *for* $U, V \in SL(n, \mathbb{Z})$.

Theorem (Kannan and Bachem [1979])

There exists an algorithm which returns the Smith Normal Form of a given nonsingular $n \times n$ *matrix* A *and the multipliers* U *and* V *and runs in time polynomial in* n *and* $\max |a_{ij}|$ *where* $A = (a_{ij})$.

General Binomial Systems

$$\left\{ \begin{array}{cccc} x^{a_1} & - & c_1 & = & 0 \\ \vdots & & \vdots & & \vdots \\ x^{a_n} & - & c_n & = & 0 \end{array} \right. \rightarrow \left\{ \begin{array}{cccc} x_1^{a_{11}} x_2^{a_{12}} \cdots x_n^{a_{1n}} & - & c_1 & = & 0 \\ \vdots & & \vdots & & \vdots \\ x_1^{a_{n1}} x_2^{a_{n2}} \cdots x_n^{a_{nn}} & - & c_n & = & 0 \end{array} \right.$$

where each $a_i \in \mathbb{Z}^n$ and $c_i \in \mathbb{C}*$, and $x = (x_1, x_2, \ldots, x_n)$.

$$\downarrow$$

$$(x_1, \ldots, x_n)^A - (c_1, \ldots, c_n)^I = 0$$

where A is the matrix of exponents and I is the identity matrix.

$$\downarrow$$

$$f(x_1, \ldots, x_n) = \left\{ \begin{array}{ccc} x_1^{s_{11}} - c_1^{v_{11}} \cdots c_n^{v_{n1}} & = & 0 \\ \vdots & \vdots & \vdots \\ x_n^{s_{nn}} - c_1^{v_{1n}} \cdots c_n^{v_{nn}} & = & 0 \end{array} \right.$$

General Binomial Systems

Algorithm for General Binomial Systems

Input: a general binomial system $f(x) := x^A - c$.

1. *Use Kannan and Bachem's algorithm to put* A *into Smith Normal Form:* $UAV = S$.
2. *Let* ε *be a suitable lower bound for* $\frac{3-\sqrt{7}}{2\gamma(f,\zeta)}$ *where* ζ *is a true root of* f
3. *Approximate the roots of the (diagonal) system* $x^S - c^V = 0$ *to within* $\frac{\varepsilon}{\sqrt{n}||U||}$ *with some* $z = (z_1, \ldots, z_n)$.
4. *Let* $\alpha = z^U$ *and return* α.

Proposition

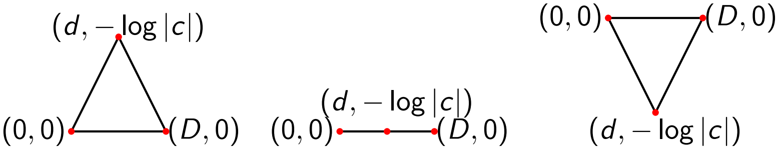
The above algorithm has average case complexity $O((n(\log d + \log n) + d)^3(\log(n(\log d + \log n) + d))^2)$.

Trinomials: $1 + cx_1^d \pm x_1^D$

Example

For $f(x_1) := 1 + cx_1^d \pm x_1^D$ *with* $c \in \mathbb{C} \setminus \{0\}$ *the lower polynomials of* f *are*

- $1 \pm x_1^D$ *if* $0 < |c| < 1$
- f *if* $|c| = 1$
- $1 + cx_1^d$ *and* $cx_1^d \pm x_1^D$ *if* $|c| > 1$



Trinomials: $1 + cx_1^d \pm x_1^D$

Definition – W-Property (Avendaño [2008])

Suppose $f(x_1) := c_1 x_1^{a_1} + \ldots + c_t x_1^{a_t} \in \mathbb{C}[x_1]$. We say f has the W-property iff the following implication holds: $(a_i, -\log|c_i|)$ is within vertical distance W of the lower hull of ArchNewt$(f) \implies (a_i, -\log|c_i|)$ is a lower vertex of ArchNewt(f).

Proposition (Avendaño [2008])

Let $f(x_1) := 1 + cx_1^d \pm x_1^D$. If f satisfies the W-property with $W \geq \log_2(36D^2)$ then any nonzero root x of a lower binomial of f satisfies $\alpha(f, x) < \alpha_0$.

Trinomials: $1 + cx_1^d \pm x_1^D$

Robust α Theorem (Blum et al. [1998])

There are positive real numbers α_0 and u_0 such that if $\alpha(f, z) < \alpha_0$, then there is a root ζ of f such that

$$B\left(\frac{u_0}{\gamma(f, z)}, z\right) \subset B\left(\frac{3 - \sqrt{7}}{2\gamma(f, \zeta)}, \zeta\right)$$

Trinomials: $1 + cx_1^d \pm x_1^D$

Algorithm for $1 + cx_1^d \pm x_1^D$

Input: $f(x_1) := 1 + cx^d \pm x^D$.

1. *If $d = 1$ and $D = 2$ use the quadratic formula to solve for the roots of f.*
2. *Otherwise if f has the W-property, use the algorithm for monic univariate binomials to approximate a root of the lower binomial of degree D to within $\frac{\varepsilon}{(3-\sqrt{7})10}$, where ε is as in the univariate binomial case.*

Lemma (B.)

On average this algorithm has computational complexity $O((\log d)^3(\log\log d)^2)$.

General Trinomials

Let $f(x_1) := c_1 + c_2 x_1^d + c_3 x_1^D$, $\mu = \frac{1}{c_1}$, $\rho = \left(\frac{c_1}{c_3}\right)^{\frac{1}{D}}$, and $\nu = \frac{c_2}{c_1}\left(\frac{c_1}{c_3}\right)^{\frac{d}{D}}$, and observe that

$$\mu f(\rho x_1) = \mu c_1 + \mu c_2 \rho^d x_1^d + \mu c_3 \rho^D x_1^D x$$

$$= 1 + \nu x_1^d \pm x^D$$

Future Work

- Handling trinomials that do not satisfy the W-property
- Systems of trinomials
- Approximating a real root or a root near a query point

References

Martín Avendaño. Unpublished notes, 2008.

Lenore Blum, Felipe Cucker, Mike Shub, and Steve Smale. *Complexity and Real Computation*. Springer-Verlag, 1998.

Ravindran Kannan and Achim Bachem. "Polynomial Algorithms for Computing the Smith and Hermite Normal Forms of an Integer Matrix". *SIAM Journal on Computing*, 8, 1979.

Steve Smale. "Newton's Method Estimates from Data at One Point". In *The Merging of Disciplines: New Directions in Pure, Applied, and Computational Mathematics*, 1986.