Solving System of Polynomial Equations

Carlos Osco Huaricapcha

Summer 2017, Mathematics Science Research Institute

Notes written from Maurice Rojas' lectures.

I'd love to hear your feedback. Feel free to email me at coscohua@mail.sfsu.edu.
See git:icarlitoss/msri-up for updates.

Contents

1 Introduction, Algebraic Geometry

← June 26, 2017

1.1 What is it?

Ideally, Algebraic Geometry occurs over rings other than $\mathbb{C}$. Let's see some examples of polynomial solving.

Example 1.1. Let $N = pq$ where N is a positive integer and $p, q > 1$. (Factoring.) Usually, we are given N, so we can find p and q.
The security of the R.S.A. Crypto-system is based on this *hardness*.

Example 1.2. Equilibria in Chemical Analysis systems (reaction networks).

$$\begin{aligned} \text{linear ordinary differential equation} \rightarrow \quad & \dot{x} = 3x + 7 \\ \text{non-linear ordinary differential equation} \rightarrow \quad & \dot{x} = \text{polynomial}(x)(\deg \geq 2) \end{aligned}$$

Differential equations govern chemical reaction rates. So, when the reaction settles down the concentrations reactions an *equilibrium*, and the concentrations wind up being *real* solutions to a system of polynomial equations. It is fair to say that *real solving* in fact makes up to more than 30% of electrical engineering.

Example 1.3. $\mathbb{F}_1$: Finite field with q elements in coding theory and cryptography. You often do arithmetic over $\mathbb{F}_q$, and polynomial system solving occurs very often. (q=prime power).

Example 1.4. $\mathbb{Q}_p$(p-acid reactions): This field is related to the solution of Weil-conjectures around the mid-20th century. Multiple fields medals were ensued.

Fun fact: 40% of the field medals are awarded to Algebraic Geometry.
Let's start with easy equations (1 equation, 1 variable, 2 terms, $\mathbb{C}$).

Example 1.5. Given $c_1, c_2 \in \mathbb{C}; a_1, a_2 \in \mathbb{Z}$. How do we solve

$$c_1 x_1^{a_1} + c_2 x_1^{a_2} = 0$$

? Notice that it is easy to reduce to the case

$$x^d = c \quad \text{where} \quad d \in \mathbb{Z}; c \in \mathbb{C}.$$

In fact if you throw-out $x_1 = 0$, then we may assume $c \neq 0$.

1.2 Euler's Formula

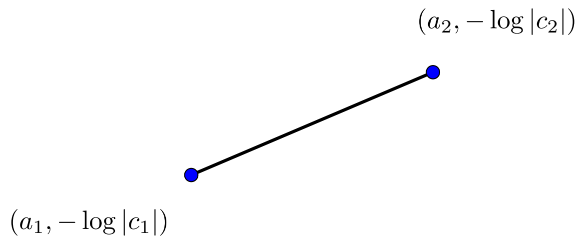
It is easy to take d$^{\text{th}}$ roots.

$$\exp^{\sqrt{-1}\tau} = \cos\tau + \sqrt{-1}\sin\tau.$$

Exercise 1. What is a simple formula for the roots of $x_1^d = c$ using Euler's formula?
Important observation: Polynomials and polytopes are long lost siblings.

1.3 ArchNewt and Newt

Definition. $ArchNewt(c_1x_1^{a_1} + c_2x_1^{a_2})$:=



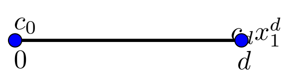
Consider the slope of this line segment(assuming $a_1 \neq a_2$).

$$\frac{\log|c_2| - (-\log|c_1|)}{a_2 - a_1} = \log|\text{all the roots of } c_1x_1^{a_1} + c_2x_1^{a_2}|\,.$$

Exercise 2. Prove it.

Here is another connection: If $f(x) = c_0 + c_1x_1 + \ldots + c_dx_1^d$. where $d \geq 1$ is the degree then f has exactly d roots counting multiplicity. (A version of the Fundamental Theorem of Algebra "F.T.A"). This fact goes back to 1609 "P. Roth," and 1629 "Argand."

Definition. $Newt(c_o + c_1x_1 + \ldots + c_dx_1^d) := [0, d]$.



<u>Observe:</u> $d = Length(Newt(\ldots))$.

1.4 Backup

1.4.1 Convex

Definition. Given any $S \subseteq \mathbb{R}^n$. We say S is convex $\iff$ for all $x, y \in S$ the line segment connecting x and y also lies in S.

$$\{\lambda x + (1-\lambda)y \mid \lambda \in [0,1]\}.$$

1.4.2 Convex Hull

Definition. Given points $a1, \dots, a_\tau \in \mathbb{R}^n$ their convex hull is

$$Conv(\{a1, \dots, a_\tau\}) := \text{smallest convex set containing} a1, \dots, a_\tau.$$

Example 1.6. $Conv(\{\dot{}\,\cdot\}) = \text{Line}$.
In $\mathbb{R}^3$ it is the wrapping ribbons around one point to another (?).

1.4.3 Polytope

Definition. A polytope (in $\mathbb{R}^n$) is just the convex hull of any finite point set in $\mathbb{R}^n$.

1.4.4 Correct definitions for Newt and Archnewt

Let's think about column vectors. If $c_1 x^{a_1} + \cdots + c_T x^{a_\tau} \quad (c_1, \dots, c_\tau \in \mathbb{C}^*)$ where $X = (x_1, \dots, x_n)$ and $X^{a_J} = x_1^{a_{1,J}}, \dots, X_n^{a_{n,J}}$. Then,

$$\begin{aligned} Newt(f) &:= \text{Conv}(\{a_1, \dots, a_\tau\}) \\ ArchNewt(f) &:= \text{Conv}(\{(a_J, -\log|c_J|) \mid \text{J} \in \{1, \dots, \tau\}\}). \end{aligned}$$

Exercise 3. Find each of following: $Newt(0), Newt(1-x_1), Newt(1-x_1+x_2+x_1x_2)$, and $ArchNewt(1+1000x_1+x_1^3)$.

<u>Note:</u>

$$\begin{aligned} \mathcal{A} &:= [a_1, \dots, a_\tau] \\ &= \begin{bmatrix} a_{1,1} & \dots & a_{1,\tau} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \dots & a_{n,\tau} \end{bmatrix} \in \mathbb{Z}^{n \times \tau} \end{aligned}$$

is called the support of f.

1.5 Univariate Trinomials

Let $f(x_1) = 1 - 1000x_1^{18} + x_1 51$. What can we say about the ($\mathbb{C}$) roots?

$$\text{“Physicists Approach…”: } |x_1| \begin{cases} \text{small} \\ \text{big} \end{cases}$$

When $|x_1|$ is *small:* $|x_1|^{18} \gg |x_1|^{51}$. Perhaps, the roots of $1 - 1000x_1^{18}$ are close to the roots of $1 - 1000x_1^{18} + x_1^{51} \Rightarrow$ maybe (?) there are <u>18</u> roots of f with norm $\sqrt[18]{\frac{1}{1000}}$ (small radius). Also, when $|x_1|$ is *big:* $|x_1|^{51} \gg |x_1|^{18} \Rightarrow$ maybe(?) the nonzero roots of $\underline{1 - 1000x_1^{18}} + x_1^{51}$ are near the nonzero roots of f. Perhaps, f has 33 roots of norm near $\sqrt[33]{1000}$ (big radius).

<u>Truth:</u> 18 roots near inner orbit and 33 in the bigger orbit.

pic goes here of the orbits

Exercise 4. How close are the norms of the roots of f to the 2 approximations?

Observe: Archnewt(f) is the following graph closed by the points $(0,0)$, $(51,0)$, and $(18, -\log 1000)$.

Graph of the triangle goes here

slope "s" of the lowets segements are close to the log |roots of f|.

2 Archimedean Tropical Variety

2.1 ArchTrop and Amoeba

The connection between polynomials and polytopes is closely related to Tropical Geometry. Historically, the roots of the connection go back to Newton around 1676, I. Newton wrote a letter describing how to extract power series solutions $y = y(x)$ to polynomial equations like $f(x, y) = 0$.
Also, J. Hadamard around 1896, derived a version of $ArcNewt$ for the power series, and observed a connection between slopes of edges and location of roots.

Definition. For one variable:

$$ArchTrop(f) = \text{slopes of lower edges of } ArchNewt(f).$$

Definition.

$$Amoeba(f) = \{\log |J| \mid J \in \mathbb{C}, f(J) = 0, J \neq 0\}.$$

Theorem (Avendaño, Kogan, Nisse, Rojas.(2013))**.** *If $f(x_1) = c_1 x^{a_1} + \cdots + c_\tau x^{a_\tau}$ then, for any $u \in Amoeba(f)$ there is a $v \in ArchTrop(f)$ with $|u - v| \leq \log 3$.*

I.e. These sets approximate each other.

Example 2.1. Given: $T = 2$(binomials). Then, $Amoeba = 1$ point and $ArchTrop = 1$ point.

Example 2.2. Given: $T = 3$(trinomials). Then, $Amoeba =$ up to the degree many points points. $1 - 1000x^{18} + x^{51}$ has ≈ 26 points (?) . Also, $ArchTrop(f) = 2$ points. (Include graph of numb line).

Exercise 5. Figure out how many points.

2.2 Vertex, Edge, and Lower Edge

2.2.1 Vertex

Definition. Consider minimizing $\alpha x + \beta y$ for $(x, y) \in P$. If this minimum is attained at a unique point v then we call v a *vertex* of P with inner normal (α, β).

2.2.2 Edge

Definition. Consider minimizing $\alpha x + \beta y$ for $(x, y) \in P$. If this minimum is attained at a proper subset of P(with $\geq$2 points) then this set is called an *edge* with inner normal (α, β). (Include graph of norms).

2.2.3 Lower Edge

Definition. If $P \in \mathbb{R}^2$ is a polygon then $E \subset P$ is a lower edge if and only if E is an edge with inner normal (α, β) for some $(\beta > 0)$. (Include graph of polygon).

3 ArchTrop and Amoeba in multivariable functions

Consider $f(x_1, x_2) = 1 + x_1 + x_2$.

Definition. If f is any polynomial in $x = (x_1, \ldots, x_n)$ then

$$Amoeba(f) := \{(\log |J|, \ldots, \log |J_n|) \mid f(J_1, \ldots J_n) = 0; J_1, \ldots J_n \in \mathbb{C}^*\}.$$

Definition.

$ArchTrop(f) := \{v \in \mathbb{R}^n \mid (v_1, -1)$ is an outer normal to a face of $ArchNewt(f)$ of positive dimension$\}$.

Note: For $n = 1$, ($ArchNewt \subset \mathbb{R}^2$) a lower edge has slope $v \iff$ it has an outer normal of the form $(v, -1)$.

put picture of the convace outter normal

Note: We will clarify faces and dimension in complete generality next time. For now, let's consider the *Amoeba* .

Enter pic of Amoeba here

Amoebae are useful because they give some intuition to begin higher dimensional zero sets.
Note: $f(x_1, x_2)$ where $x_1, x_2 \in \mathbb{C}$ has 2 manifold in $\mathbb{R}^+$.
Let's analyze why is that picture true.

$$\text{If } 1 + x_1 + x_2 = 0 \text{ then } \begin{cases} x_1 + x_2 = -1 & \Rightarrow |x_1 + x_2| = 1 & (i) \\ x_1 = -1 - x_2 & \Rightarrow |x_1| = |1 - x_2| & (ii) \\ x_2 = -1 - x_1 & \Rightarrow |x_2| = |1 + x_1| & (iii) \end{cases}$$

So $u := |x_1|, v := |x_2|$ implies that

$$\begin{aligned} 1 &\leq u + v \\ u &\leq 1 + v \\ v &\leq 1 + u. \end{aligned}$$

So $(u, v) \in$

Show graph of the rect in diag

then by taking the log of the figure, we would obtain the $Amoeba(1+x_1+x_2) = (\log |J_1|, \log |J_2|)$.

Show graph of Amoeba

Q&A: What is the relation of the product fg with $Newt(fg)$ and $ArchNewt(fg)$.

$$Newt(fg) = Newt(f) + Newt(g)$$
$$ArchNewt(fg) = \ldots \text{::: mess :::} \ldots.$$

It is a mess because $ArchNewt((1+x)^3) \neq 3ArchNewt(1+x)$.

3.1 Orientation

← June 27, 2017

Definition. An *orientation* of a line $L(\subset \mathbb{R}^2)$ is just attaching a nonzero vector $v(\in \mathbb{R}^2)$ to L.

line x+y-5

Note: This allow us to speak of *left* and *right*.

3.2 Puzzle

Definition. Given an oriented line L(through $\overrightarrow{o}$) and a point P_1.

How do we determine if P belongs to left of L or the right L? $L = x-$axis, orientation vector $v = (1, 0)$.

Ray oriented pic goes here

Example 3.1. What side is P on? (Generalize).
Sign of the determinant would determine if it is right or left.

$$\text{if } \det[v,p] \begin{cases} > 0 & \iff p \text{ is left of } L. \\ = 0 & \iff p \in L. \\ < 0 & \iff p \text{ is right of } L. \end{cases}$$

4 An Incremental Algorithm

To complete $Conv$ in $\mathbb{R}^2$ is a basic problem in computational Geometry. Here is a simple way to do it.
Input: $P_1, \ldots, P_N \in \mathbb{R}^2$. Output: A list of indices $\{c_1, \ldots, c_k\}$. such that:
(0) : the vertices $Conv\{P_1, \ldots, P_N\}$ are exactly $P_i, \ldots, P_{i_k}$.
(1) : $P_i, \ldots, P_k$ are ordered $C.C.W.$

4.1 Description

- (-1) : If $P_1 = \ldots = P_N$ then $Conv = P_1$ and you are done ☺.

- (0) : If $\det[P_2 - P_1, P_i - P_1] = 0$ (Assume $P_2 - P_1 \neq 0$) for all $i \geq 3$ then let P_{i_1} be any point.
  - minimizing: $P_{i_1} \cdot (P_2 - P_1)$, then P_{i_2} be any point.

- maximizing: $P_{i_2} \cdot (P_2 - P_1)$, then $Conv$ = the line segment connecting P_{i_1} and P_{i_2}.

- (1) : Find 3 points $\overbrace{P_1, P_2, P_3}^{W \log}$ with $\det(P_2 - P_1, P_3 - P_1) \neq 0$. i.e. Forming a non-degen triangle.
- (2) : Let $S = \{1, 2, 3\}$. Let $J := 4$
  (a) Take P_J and check if $P_J \in Conv$ defined by S.
  (b) If not find the uniques edges $\overline{P_J - P}$ and $\overline{P_J P_v}$ of $Conv((\bigcup_{l \in S} P_l) \cup P_J)$
- (3) : Update $S, J, J = J + 1$. Go to (2) unless no points left.
- (4) : Output S.

Example 4.1. Let's go back to the beginning, and trace the convex hull.

Graph of the progress goes here

4.2 Complexity

I.e. how much work? and how do you measure?
Input size: N(# of points).
Work: # of Field operations $(+, -, \times, \div)$ and sign checks involving $\mathbb{R}$ numbers.
Measuring this way, we can see that steps in (-1) and (0) use:

(i) $\leq 2(N - 1) =$ checks.

(ii) ≤ 2 subtractions of vectors.

(iii) $N - 1$ determinants and sign checks.

Hence, it takes about $2 + 3(N = 1)$ operations in the end. $\Rightarrow O(\mathcal{N})$ops.
The remaining steps then involve:

(i) 3 left and right checks, update list.

(ii) <4 left and right checks. $\vdots$

(iii) $N - 1$ left and right checks.

Hence,$\Rightarrow O(\mathcal{N}^2)$ operations in the end. $\Rightarrow O(\mathcal{N}^2)$ops . Therefore, it implies that $O(\mathcal{N})$ complexity required.
This algorithm is from the 1980's and, while not the fastest, it is conceptually one of the easiest. Also easy to extend to arbitrary dimensions.

<u>Note:</u> There are $O(\mathcal{N} \log \mathcal{N})$ convex hull algors (in $\mathbb{R}^2$). It turns out that every algorithm (for $Conv \in \mathbb{R}^2$) has worst-case complexity $\Omega(\mathcal{N} \log \mathcal{N})$.
In dimension d, Chazelle and Edelsbrunner proved in the 1980's that $Conv(N\text{-points})$ admits a $O(N^{\lfloor d/2 \rfloor} + N \log N)$ algorithm, and this is optimal.

4.3 Alternative Definition for ArchTrop(f)

Definition. If $f(x) = c_1x^{a_1} + \ldots + c_\tau x^{a_\tau}$.

$$ArchTrop(f) : \left\{ v \in \mathbb{R}^n : \max_{J \in \{1, \ldots, \tau\}} |\ c_J e^{a_J \cdot v}\ | \ \text{ is attained at } \ \geq 2 \text{ distinct points} \right\}.$$

<u>Note:</u> Tropical Geometry is nothing more than checking when 2 things agree.
Exercise 1. Binomial case in n variables. Check that definition via $ArchNewt$ = definition via max.

Example 4.2. Given: $1 - 1000x_1^{18} + x_1^{51}$. There are 3 terms. Then 3 possibilities.

(i) $|1| = |-1000e^{18v}| \quad > |e^{51v}|$.

(ii) $|-1000e^{18v}| = |e^{51v}| \quad > |1|$

(iii) $|e^{51v}| = 1 \quad |-1000e^{18v}$.

Then, solving each case we would get that

(i) $v = \frac{1}{18} \log 1000$ (and e^{51v} is simply smaller).

(ii) $v = \frac{1}{33} \log 1000$ (and$e^{18v} > 1$.

(iii) $v = 0$ (but $1 > 1000$), therefore we exclude it.

4.4 Back up

4.4.1 Affine

Definition. An *affine* subspace S of $\mathbb{R}^n$ is just a translate of a (linear) subspace of $\mathbb{R}^n$.

4.4.2 Face

Definition. A *face* Q of a polyhedron $P \subseteq \mathbb{R}^n$ is just equivalently:

(a) $P \bigcap$ supporting hyperplane.

(b) $\{x \in P | v \cdot x \text{ is minimal}\}$ for some inner normal v.

(c) $\{x \in P | w \cdot x \text{ is maximall}\}$ for some outerr normal w.

4.4.3 Half Space

Definition. A *half space* $\subset \mathbb{R}^n$ is just $H = \{x | x \cdot x \leq c\}$ for some v,

4.4.4 Dimension

Definition. The *dimension* of Q is just the dimension of the smallest affine subspace.