A Polyhedral Homotopy Continuation Method for Computing All Solutions of a Polynomial System of Equations in Complex Variables

Masakazu Kojima, Yang Dai, Sunyoung Kim, and Akiko Takeda

October 2001
A polyhedral homotopy continuation method for computing all solutions of a polynomial system of equations in complex variables

Masakazu Kojima, Yang Dai, Sunyoung Kim and Akiko Takeda

Contents

1. A polynomial equation system
2. Typical benchmark polynomial systems
3. Rough sketch of the polyhedral homotopy method
4. Basic idea of Phases 1 and 2 of the polyhedral homotopy method
5. Phase 1 — Construction of a family of homotopy functions
6. Numerical results on Phase 1
7. Phase 2 — Tracking homotopy paths
8. Numerical results on Phase 2
9. Summary

solution of a polynomial system of equations in complex variables

A polyhedral homotopy continuation method for computing all solutions of a polynomial system of equations in complex variables
Find all isolated solutions in $C^n$. Let $z \in C^n$.

\[
\begin{align*}
\bar{z}x^2x^1x(\bar{z} + 3) &= (\bar{z}x^1x)^3f \\
\bar{z}x^1x - \bar{z}x^2x - zx^1x^2 &= (\bar{z}x^2x^1x)^2f \\
\bar{z} + \bar{z}x^2x(\bar{z} + 1) - z &= (\bar{z}x^2x^1x)^2f \\
'((x)\bar{z}f, (x)zf, (x)f) &= (x)f, (\bar{z}x^2x^1x)f = x, \bar{z} = u
\end{align*}
\]

Example

\[
\begin{align*}
\cdots x^2x \cdots \cdots x^1x \cdots x^2x^1x = (x)f \\
\cdots'((x)^uf, \cdots, (x)^zf, (x)f) &= (x)f \\
'\subseteq (x^2x^1x^2x^1x) &= x
\end{align*}
\]

Where

\[
0 = (x)f
\]

I. A polynomial equation system
Find all isolated solutions in $\mathbb{C}^n$.

Here we assume that both objective and constraint functions are polynomials.

If we compute all the Karush-Kuhn-Tucker stationary solutions, the we can pick up a global optimal solution among them.

If we compute all the Karush-Kuhn-Tucker stationary solutions, we can find a global optimal solution.

Global optimization

Various engineering applications

A fundamental problem in numerical mathematics
\[ z - u^2 \quad u \]

\[ \ldots \]

96 4 096
8 2048
7 1024
6 512
5 256

\( z \) of isolated solutions

\[ 0 = 1 + \frac{1}{u}x + \cdots + \frac{1}{u}x + \frac{1}{u}x \]
\[ 0 = (1 - u) - \frac{1}{u}x^{1-u}x \]
\[ 0 = \left( z - u \right) - \frac{1}{u}x^{1-u}x + \frac{1}{u}x^{1-u}x \]
\[ \ldots \]
\[ 0 = \frac{1}{u}x^{1-u}x \left( \frac{1}{u}x^{1-u}x + \cdots + \frac{1}{u}x^{1-u}x + \frac{1}{u}x \right) \]
\[ 0 = 1 - \frac{1}{u}x \left( \frac{1}{u}x^{1-u}x + \cdots + \frac{1}{u}x^{1-u}x + \frac{1}{u}x \right) \]

2. Typical benchmark test problem — Economic polynomial
Parallel computation

Enormous computational power for solving large scale problems

Typical benchmark test problem — 2: Cyclic polynomial

\[ f(x) = x^1 + x^2 + \cdots + x^n \]

\[ f(x) = x^{2n} \]

\[ \begin{array}{cccc}
30,624 & 15,312 & 367,488 & 18,796 \\
16,796 & 8,398 & 94,796 & 11,490 \\
3,494 & 1,447 & 34,940 & 10,494 \\
\end{array} \]

\# of nonsingular isolated solutions

\[ (u(n)/2, n) \]

\[ \begin{array}{cccc}
\ldots & \ldots & \ldots & \ldots \\
1 - u x^{1-u} x \cdots z x^{1} x & = (x)^{u} f \\
1 - u x \cdots z x^{1} x & = (x)^{1-u} f \\
1 - u x \cdots z x^{1} x & = (x)^{z-u} f \\
\ldots \\
1 - u x \cdots z x^{1} x & = (x)^{z} f \\
u x + \cdots + z x + 1 x & = (x)^{1} f \\
\end{array} \]
Phase 1. Construct a family of homotopy functions.

Phase 2. Trace homotopy paths by predictor-corrector methods.

Highly nonlinear homotopy paths that require sophisticated step control.

Suitable for parallel computation; all solutions can be computed independently in parallel.

Currently the most powerful and practical method for computing all solutions of a polynomial equation system.

- Nonlinear combinatorial optimization problems.
- Large scale linear programs.
- Branch-and-bound methods.
- Branch-and-bound methos.

Rough sketch of the polyhedral homotopy (continuation) method.
Phase 3. Verify that all isolated solutions are computed.

- Approximate solutions are computed but exact solutions are never computed.
- The number of solutions is unknown in general.
4. Basic ideas of Phases 1 and 2.

Phase 1. Let $x^*$ be a solution of $f(x) = 0$. We construct a homotopy equation system $h(x, t) = 0$ such that (i) all solutions of the initial system $h(x, 0) = 0$ are known, (ii) $h(x, 1) = f(x)$ for every $x \in \mathbb{C}^n$; hence if $h(x, 1) = 0$, $x$ is a solution of $f(x) = 0$, and (iii) $x^*$ is connected to a solution $x^1$ of $h(x, 0) = 0$ through the solution path of $h(x, t) = 0$.

Phase 2. Starting from each known solution of the initial system $h(x, 0) = 0$, we trace the solution path of $h(x, t) = 0$ till $t$ attains 1 by a predictor-corrector method to obtain a solution of $f(x) = 0$. 
This idea is common for the traditional linear homotopy method.

- Some solution paths diverge as \( t \to \infty \) and tracing such paths are useless.
- The number of useless divergent paths is much less in the polyhedral method.
- Multiple homotopy functions are employed in polyhedral homotopy methods while a common single \( y \) is employed for all solutions in the traditional linear homotopy method.

\[
\text{Given } f(x) = 0 \text{ in the traditional linear homotopy method,}
\]
We call \( \mathcal{A} \) the support of \( f \) \( u \mathbb{Z} \) the subset of \( \mathbb{Z} \) and define \( x_a = x_{a_1} \cdots x_{a_n} \),

\[
\mathcal{A}(x)_f = \bigcup \{ a : 0 \preceq (a, \ldots, a) \} \equiv \mathcal{Z} \mathcal{A} \nabla \mathcal{A}
\]

as

\[
((x)_u f, \ldots, (x)_1 f) = (x)_f \text{ of a poly system } \Rightarrow \mathcal{A}(x)_f \mathcal{A} \nabla \text{ of a poly system } \Rightarrow (x)_f \mathcal{A} \nabla \text{ poly system } \Rightarrow \mathcal{A}(x)_f \mathcal{A} \nabla \text{ poly system }
\]

\[
\begin{align*}
\mathcal{Z} \mathcal{A} \nabla \mathcal{A} &= \mathcal{A}(x)_f \mathcal{A} \nabla \text{ poly system } \\
(x)_u f \cdots (x)_1 f &= (x)_f \text{ poly system }
\end{align*}
\]

\[
\left\{ u \mathbb{Z} \in \mathbb{Z} : 0 \preceq (a, \ldots, a) \right\} \equiv \mathcal{Z} \mathcal{A} \nabla \mathcal{A}
\]
\[
\begin{align*}
\alpha &= (\beta, \gamma, \delta, \varepsilon) \\
\beta &= (\alpha, \beta, \gamma, \delta) \\
\gamma &= (\alpha, \beta, \gamma, \delta) \\
\delta &= (\alpha, \beta, \gamma, \delta) \\
\varepsilon &= (\alpha, \beta, \gamma, \delta) \\
\chi &= (\alpha, \beta, \gamma, \delta) \\
\end{align*}
\]

For example, \( \varepsilon = \gamma \),

where

\[
\begin{align*}
\varphi \in (\alpha, \beta, \gamma, \delta) \quad &\text{such that} \\
\psi \in (\alpha, \beta, \gamma, \delta) \\
\end{align*}
\]

\[
\begin{align*}
\phi \in \varphi &
\Rightarrow \\
\psi \in \psi &
\Rightarrow \\
\end{align*}
\]
Phase I. Construction of a family of homotopy functions

\[ h_k(x(t)) = \begin{cases} 
0 & \text{if } k \neq 1, 2, \ldots, n \\
1 & \text{if } k = 1, 2, \ldots, n 
\end{cases} \]

where exactly two of \( f^\alpha \) are positive and all others are zero. For \( k = 1, 2, \ldots, n \)

\[ (u, \ldots, z = \gamma) \quad (\alpha) \quad \forall \alpha \in \{ a \in A \} \quad \therefore x(x) \quad \therefore \gamma \quad \text{for} \quad A \quad \text{for} \quad (a) \]

\[ (u, \ldots, z = \gamma) \quad [0, 1] \times w \in C \quad \exists \quad (t, x) \quad (t, x) \quad \gamma \]

\[ (u, \ldots, z = \gamma) \quad \text{for} \quad (a) \quad \exists \quad (t, x) \quad (t, x) \quad \gamma \]
Each component $h_{k_j}(x)$ consists of two terms; hence the starting equation system turns out to be a binomial equation system.

We can easily compute all solutions by linear algebra (or elimination technique).

\begin{align*}
(n' = 1, \ldots, 2)
0 &= \sum_{j=1}^{n'} \left( \frac{x}{\mu} x_{j' \mu} \right) c_j + \sum_{j=1}^{n'} \left( \frac{x}{\mu} x_{j' \mu} \right) c_j \\
&\equiv (0, x)_{\mu} y
\end{align*}
(b) ∀ sol. \( x^* \) of \( f(x) = 0 \), \( \exists k \), \( \exists \) sol. \( \tilde{x} \) of \( h^k(x, 0) = 0 \); \( \tilde{x} \) is connected to \( x^* \) through a sol. path \( C = \{(\xi(t), t) : t \in [0, 1]\} \) of \( h^k(x, t) = 0 \).
How do we construct such a family of homotopy functions?

Let \( \mathbf{v} = (v_{1},\ldots,v_{n}) \in \mathbb{R}^{n} \), randomly (randomly from \( \mathbb{R}^{n} \)).

Choose \( m \) \( \in \mathbb{Z} \)

\( A = (a_{1},\ldots,a_{n}) \)

whose value we will determine later. Define

\[
\forall \mathbf{v} \in \mathbb{R}^{n} \quad \exists \mathbf{v} \in \mathbb{R}^{n} \quad \forall \mathbf{v} \in \mathbb{R}^{n} \quad \forall \mathbf{v} \in \mathbb{R}^{n} \quad \forall \mathbf{v} \in \mathbb{R}^{n} \quad \forall \mathbf{v} \in \mathbb{R}^{n} \quad \forall \mathbf{v} \in \mathbb{R}^{n}
\]

If

\[
\forall \mathbf{v} \in \mathbb{R}^{n} \quad \exists \mathbf{v} \in \mathbb{R}^{n} \quad \forall \mathbf{v} \in \mathbb{R}^{n} \quad \forall \mathbf{v} \in \mathbb{R}^{n} \quad \forall \mathbf{v} \in \mathbb{R}^{n} \quad \forall \mathbf{v} \in \mathbb{R}^{n} \quad \forall \mathbf{v} \in \mathbb{R}^{n}
\]

Then

\[
\forall \mathbf{v} \in \mathbb{R}^{n} \quad \exists \mathbf{v} \in \mathbb{R}^{n} \quad \forall \mathbf{v} \in \mathbb{R}^{n} \quad \forall \mathbf{v} \in \mathbb{R}^{n} \quad \forall \mathbf{v} \in \mathbb{R}^{n} \quad \forall \mathbf{v} \in \mathbb{R}^{n} \quad \forall \mathbf{v} \in \mathbb{R}^{n}
\]

is at most \( 2n \) equations.

Nondegeneracy assumption: A solution of (1), \( \forall \mathbf{v} \in \mathbb{R}^{n} \), exists for exactly 2 of \( \mathbf{v} \). Let \( \mathbf{v} \in \mathbb{R}^{n} \), let

\[
(2) 
\begin{align*}
\forall \mathbf{v} \in \mathbb{R}^{n} \quad & \mathbf{v} \in \mathbb{R}^{n} \\
\forall \mathbf{v} \in \mathbb{R}^{n} \quad & \mathbf{v} \in \mathbb{R}^{n} \\
\forall \mathbf{v} \in \mathbb{R}^{n} \quad & \mathbf{v} \in \mathbb{R}^{n} \\
\forall \mathbf{v} \in \mathbb{R}^{n} \quad & \mathbf{v} \in \mathbb{R}^{n} \\
\forall \mathbf{v} \in \mathbb{R}^{n} \quad & \mathbf{v} \in \mathbb{R}^{n} \\
\forall \mathbf{v} \in \mathbb{R}^{n} \quad & \mathbf{v} \in \mathbb{R}^{n} \\
\forall \mathbf{v} \in \mathbb{R}^{n} \quad & \mathbf{v} \in \mathbb{R}^{n}
\end{align*}
\]

where

\[
\begin{align*}
(2') \quad & \mathbf{v} \in \mathbb{R}^{n} \\
(2') \quad & \mathbf{v} \in \mathbb{R}^{n} \\
(2') \quad & \mathbf{v} \in \mathbb{R}^{n} \\
(2') \quad & \mathbf{v} \in \mathbb{R}^{n} \\
(2') \quad & \mathbf{v} \in \mathbb{R}^{n} \\
(2') \quad & \mathbf{v} \in \mathbb{R}^{n} \\
(2') \quad & \mathbf{v} \in \mathbb{R}^{n} \\
(2') \quad & \mathbf{v} \in \mathbb{R}^{n}
\end{align*}
\]

In the polyhedral homotopy theory, it is known that

\[
\{[0,1] \times \exists t : (t', (t, x))\} = \mathcal{C} \text{ path through } a \text{ sol. } x \neq x \text{ is connected to } x \text{ through a } x. \\
0 = (0, x)_{\mathcal{C}, x} \text{ sol of } x \in \mathcal{I} \in (\mathcal{E}', \mathcal{A}) \in 0', 0 = (x)f \text{ sol of } A \cup (b) \\
(\mathcal{I} \in (\mathcal{E}', \mathcal{A})) \text{ satisfies the desired properties we have mentioned.}
\]

\{\mathcal{I} \in (\mathcal{E}', \mathcal{A})\} \equiv (2) \text{ is finite.} \\
\text{The family } y \text { satisfies } (\mathcal{I} \in (\mathcal{E}', \mathcal{A})) (t', x)_{\mathcal{E}', \mathcal{A}} \text{ solutions of } x \text{ is finite.} \\
\text{In the polyhedral homotopy theory, it is known that}
Therefore computing all solutions of the linear inequality system (1) forms an important subprob. in Phase 1. Parallel computation.

- Implicit enum. tech. (or b-and-b. methods) used in optimization.
- The simplex method for linear programs.

↓

with the comb. cond. (2)
DEC Alpha 21164 (600 MHz) with 1 GB memory

6. Computational results on the solution of (1) $\vec{x}$ (2) $\overline{1}$
Parallel Comp. on the sol. of (1) & (2) — Eco-n Problems

Intel Pentium III (824MHz) with 640MB memory
Parallel Comp. on the sol. of (1) $\mathcal{C}$ (2) $\mathcal{C}_n$ problems

Intel Pentium III (824MHZ) with 640MB memory
Homotopy equation system

Starting from a known initial solution \((x_0, 0, 0)\), trace the solution path.

\[
\begin{align*}
(\forall \alpha \in 0 \leq \tau \leq 1) & \\
\forall \alpha \in 0 \leq \tau \leq 1 & \\
(\exists) & \\
[0, 1] \times w \in (\tau, x) & \\
0 & = (D)_{\tau} \int_{D} x(\alpha) \, c \times \gamma \equiv (\tau, x) \, \gamma
\end{align*}
\]
\[ \dot{x} = x - x_0 = (\mathcal{P} + 0, 0) = 0 \]

Correct Newton method to \( h \) with the initial point \( x_0 \):

\[ 0 = \mathcal{P}(0, 0) + x(0) \cdot \mathcal{D} + x(0) \cdot \mathcal{D} : 0 < \mathcal{P} \]

Preceded with a step length \( \Delta \).
Step length control is essential!

Too small step length $\rightarrow$ predictor with more predicted iter. and more CPU time.

Too large step length $\rightarrow$ jump into a different solution path.

$0 = P(\dot{y}, x) \mu_D + P(\dot{y}, x)x \mu_D : (\gamma \tau, y, x)$ at $0 < \mu$
Difficulty in Phase 2: High nonlinearity in $y(t)$, $x(t)$ is huge, for example.

Construct homotopies with less power.

Sophisticated step length control.

$$
\cdots + \int_{000}^{1} \int_{0}^{1} x'(v) f c + \int_{000}^{1} \int_{0}^{1} x'(v) f c + \int_{0}^{1} x'(v) f c + \cdots = (t, x)f(y)$$
Change of $t_p$ as $t \rightarrow 1$, $p = 10, 1,000, 10,000$
8. Numerical results — Economic problems
Numerical results — Cyclic n problems
9. Concluding Remarks

(a) While we trace a homotopy path numerically, a jump into another path sometime occurs $\iff$ Not 100% reliable. But the reliability is very high; for example, less than 0.1% solutions are missing in our numerical experiments. There are two ways to compensate such a fault.

(a-1) Suppose that numerical tracing of two paths led to a common solution as in case 1 below. Then we know there is an illegal jump while tracing one of them. Hence, recompute those two paths using smaller predictor step lengths.
Construct multiple sets of homotopy functions, each of which theoretically covers all solutions. Then apply the polyhedral homology method to each set of homotopy functions to generate multiple sets of solutions. Therefore, merging all these sets of solutions increases the reliability much. Even if a solution is missing in one set, the same solution is unlikely to be missed in all other sets. Therefore, merging all these sets of solutions increases the reliability much.
The polyhedral homotopy continuation method involves various optimization techniques such as branch-and-bound methods, linear optimization techniques, and predictor-corrector methods. This problem can be formulated as a nonlinear combinatorial optimization problem. Reducing the powers of the continuation parameter is crucial to achieve the numerical stability and efficiency in tracking homotopy paths. An important feature of the homotopy continuation method is that all homotopy paths can be computed independently and simultaneously in parallel.

Concluding Remarks