Solving multivariate polynomial systems (Assignment 2)

Chuen Long Mountain Chan / B1-2

December 18, 2019

In this paper, we will develop mathematical theories and algorithms for the following problem.

In: $f \in \mathbb{C}[x_1, \ldots, x_n]^n$

Out: $S \subset \mathbb{C}^n$, the set of complex solutions of f.

1 Solving by Multiplication Matrix (Pham systems)

1.1 Theory

Definition 1 (Pham system) *A system $f \in \mathbb{C}[x_1, \ldots, x_n]^n$ is called a Pham system if it has the following form:*

$$
\begin{aligned}
f_1 &= x_1^{d_1} + \sum_{e_1+\cdots+e_n<d_1} a_{1,e} x_1^{e_1} \cdots x_n^{e_n} \\
&\vdots \\
f_n &= x_n^{d_n} + \sum_{e_1+\cdots+e_n<d_n} a_{n,e} x_1^{e_1} \cdots x_n^{e_n}
\end{aligned}
$$

Definition 2 (Monomial Basis) *The monomial basis for $(d_1, \ldots, d_n)$ is given by*

$$\omega = [x_1^{e_1} \cdots x_n^{e_n} : 0 \leq e_1 < d_1, \ldots, 0 \leq e_n < d_n]$$

They are ordered in the decreasing order in the total degree where $x_n > x_{n-1} > \cdots > x_1$.

Definition 3 (Multiplication matrix) *Let $g \in \mathbb{C}[x_1, \ldots, x_n]$. Then the multiplication matrix M_g for g is defined by*

$$g \begin{bmatrix} \omega_1 \\ \vdots \\ \omega_m \end{bmatrix} \equiv_f M_g \begin{bmatrix} \omega_1 \\ \vdots \\ \omega_m \end{bmatrix}$$

Theorem 4 *Let $g \in \mathbb{C}[x_1, \ldots, x_n]$ be random. Let V be the set of all the eigenvectors of M_g. Then, with probability one, we have*

$$S = \{[v_{m-1}/v_m, \ldots, v_{m-n}/v_m] : v \in V\}$$

Proof. Let $f(z) = 0$, and by Theorem 9 in project 1, we know that $\{\begin{bmatrix}\omega_1\\ \vdots \\ \omega_m\end{bmatrix} : f(z_i) = 0\}$ spans M_g. Given that g is random, for all $i \neq j$, $g(z_i) \neq g(z_j)$, then we get that $\begin{bmatrix}v_1\\ \vdots \\ v_m\end{bmatrix} = v_m \begin{bmatrix}\omega_1(z_i)\\ \vdots \\ \omega_m(z_i)\end{bmatrix}$ Observe that $x_n > x_{n-1} > \cdots > x_1$, $\omega_m = 1$, thus we get $\begin{bmatrix}v_1/v_m\\ \vdots \\ 1\end{bmatrix} = v_m \begin{bmatrix}\omega_1(z_i)\\ \vdots \\ \omega_m(z_i)\end{bmatrix}$ Hence with probability one, we get $S = \{[v_{m-1}/v_m, \ldots, v_{m-n}/v_m] : v \in V\}$ ■

1.2 Algorithms

Algorithm 5 (MulMat)

In: $f \in \mathbb{C}[x_1, \ldots, x_n]^n$ *a Pham system*, $g \in \mathbb{C}[x_1, \ldots, x_n]$

Out: M,*the multiplication matrix of* g *modulo* f

1. $\omega =$ *Monomial Basis of function* f
2. $n =$ *number of omegas*
3. $M =$ *Remainder(*$g\omega_i$, ω_j*) with size* $n \times n$
4. *Return* M

Algorithm 6 (SolveByMulMat)

In: $f \in \mathbb{C}[x_1, \ldots, x_n]^n$ *a Pham system*, $g \in \mathbb{C}[x_1, \ldots, x_n]$ *random linear*

Out: $S \subset \mathbb{C}^n$, *the set of complex solutions of* f.

1. $M = MultMat(fs, vs, g)$
2. $E = Eigenvalue(M)$
3. $S = E[i-k][j]/E[i][j]$ *with* $k = 1 \ldots n$
4. *Return* S

2 Solving by Univariate Resultant (General systems)

2.1 Theory

Definition 7 (Sylvester matrix) *Let* $f = \sum_{i=0}^{m} a_i x^i, g = \sum_{i=0}^{n} b_i x^i \in \mathbb{C}[x]$, *where* $a_m = b_n \neq 0$. *Then the Sylvester matrix* S *of* f, g *is defined as*

$$S = \begin{bmatrix} a_m & a_{m-1} & \cdots & \cdots & a_0 & & \\ & \ddots & \ddots & & & \ddots & \\ & & a_m & a_{m-1} & \cdots & \cdots & a_0 \\ b_n & b_{n-1} & \cdots & b_0 & & & \\ & \ddots & \ddots & & \ddots & & \\ & & \ddots & \ddots & & \ddots & \\ & & & b_n & b_{n-1} & \cdots & b_0 \end{bmatrix}$$

where there are n rows of the coefficients of f and m rows of the coefficients of g.

Definition 8 (Subresultant) *The k-th subresultant of f and g with respect to x, written as $R_{x,k}(f,g)$ is defined as*

$$R_{x,k}(f,g) = \sum_{i=0}^{k} |S_{k,i}|\, x^i$$

where $S_{k,i}$ is the submatrix of the Sylvester matrix S, consisting of

1. *the first $n-k$ rows of the coefficients of f*
2. *the first $m-k$ rows of the coefficients of g*
3. *the first $n+m-2k-1$ columns and the $n+m-k-i$-th column*

Theorem 9 *Let $f = a_m(x-\alpha_1)\cdots(x-\alpha_m)$ and $g = b_n(x-\beta_1)\cdots(x-\beta_n)$. Then we have*

$$R_{x,0}(f,g) = a_m^n b_n^m \prod_{i,j} (\alpha_i - \beta_j)$$

Proof. For the sake of simple presentation of the proof idea, let us consider $\deg f = 3$ and $\deg g = 2$. The idea can be easily generalize to arbitrary degrees.
Let $|V| \neq 0$, then we get the Sylvester matrix

$$\begin{bmatrix} a_3 & a_2 & a_1 & a_0 & 0 \\ 0 & a_3 & a_2 & a_1 & 0 \\ b_2 & b_1 & b_0 & 0 & 0 \\ 0 & b_2 & b_1 & b_0 & 0 \\ 0 & 0 & b_2 & b_1 & b_0 \end{bmatrix}$$

then from class, we are given the matrix

$$\begin{bmatrix} a_3 & a_2 & a_1 & a_0 & 0 \\ 0 & a_3 & a_2 & a_1 & 0 \\ b_2 & b_1 & b_0 & 0 & 0 \\ 0 & b_2 & b_1 & b_0 & 0 \\ 0 & 0 & b_2 & b_1 & b_0 \end{bmatrix} \begin{bmatrix} 1 & & a_1^4 & \dots & a_3^4 \\ & 1 & a_1^3 & \dots & a_3^3 \\ & & a_1^2 & \dots & a_3^2 \\ & & a_1^1 & \dots & a_3^1 \\ & & a_1^0 & \dots & a_3^0 \end{bmatrix} = \begin{bmatrix} a_3 & a_2 & \alpha_1^1 f(\alpha_1) & \dots & \alpha_3^0 f(\alpha_3) \\ 0 & a_3 & \alpha_1^0 f(\alpha_1) & \dots & \alpha_3^0 f(\alpha_3) \\ b_2 & b_1 & \alpha_1^2 g(\alpha_1) & \dots & \alpha_3^2 g(\alpha_3) \\ 0 & b_2 & \alpha_1^1 g(\alpha_1) & \dots & \alpha_3^1 g(\alpha_3) \\ 0 & 0 & \alpha_1^0 g(\alpha_1) & \dots & \alpha_3^0 g(\alpha_3) \end{bmatrix}$$

Since α is the roots of f, we have

$$\begin{bmatrix} a_3 & a_2 & 0 & \dots & 0 \\ 0 & a_3 & 0 & \dots & 0 \\ b_2 & b_1 & \alpha_1^2 g(\alpha_1) & \dots & \alpha_3^2 g(\alpha_3) \\ 0 & b_2 & \alpha_1^1 g(\alpha_1) & \dots & \alpha_3^1 g(\alpha_3) \\ 0 & 0 & \alpha_1^0 g(\alpha_1) & \dots & \alpha_3^0 g(\alpha_3) \end{bmatrix}$$

thus

$$|\mathrm{LHS}| = |S||V|$$

and

$$|\mathrm{RHS}| = \alpha_3^2 \begin{bmatrix} \alpha_1^2 g(\alpha_1) & \dots & \alpha_3^2 g(\alpha_3) \\ \alpha_1^1 g(\alpha_1) & \dots & \alpha_3^1 g(\alpha_3) \\ \alpha_1^0 g(\alpha_1) & \dots & \alpha_3^0 g(\alpha_3) \end{bmatrix} \cdot |V| = a_3^2 g(\alpha_1)\dots g(\alpha_3) \begin{bmatrix} \alpha_1^2 & \dots & \alpha_3^2 \\ \alpha_1^1 & \dots & \alpha_3^1 \\ \alpha_1^0 & \dots & \alpha_3^0 \end{bmatrix} = a_3^2 g(\alpha_1)\dots g(\alpha_3)|V|$$

. Hence

$$|S| = a_3^2 b_2 (\alpha_1 - \beta_1)\dots(\alpha_1 - \beta_2)$$

.

$$\vdots$$

$$\begin{aligned}&b_2(\alpha_3-\beta_1)\dots(\alpha_3-\beta_2)\\&=a_3^2b_2^3\prod_{i,j}(\alpha_i-\beta_j)\end{aligned}$$

■

Theorem 10 $R_{x,k} \in \langle f, g\rangle$.

Proof. For the sake of simple presentation of the proof idea, let us consider $\deg f = 3$ and $\deg g = 2$ and $k = 1$. Consider the Sylvester matrix

$$\begin{bmatrix} a_3 & a_2 & a_1 & a_0 & 0 \\ 0 & a_3 & a_2 & a_1 & 0 \\ b_2 & b_1 & b_0 & 0 & 0 \\ 0 & b_2 & b_1 & b_0 & 0 \\ 0 & 0 & b_2 & b_1 & b_0 \end{bmatrix}$$

Using the sub-resultant elimination method we get the matrix

$$\begin{bmatrix} a_3 & a_2 & a_1 \\ b_2 & b_1 & b_0 \\ 0 & b_2 & b_1 \end{bmatrix} x^1 + \begin{bmatrix} a_3 & a_2 & a_0 \\ b_2 & b_1 & 0 \\ 0 & b_2 & b_1 \end{bmatrix} x^0 = \begin{bmatrix} a_3 & a_2 & a_1x^1 + a_0x^0 \\ b_2 & b_1 & b_0x^1 \\ 0 & b_2 & b_1x^1 + b_0x^0 \end{bmatrix} = \begin{bmatrix} a_3 & a_2 & a_3x^3 + a_2x^2 + a_1x^1 + a_0x^0 \\ b_2 & b_1 & b_2x^3 + b_1x^2 + b_0x^1 \\ 0 & b_2 & b_2x^2 + b_1x^1 + b_0x^0 \end{bmatrix}$$

$$= \begin{bmatrix} a_3 & a_2 & 1 \\ b_2 & b_1 & 0 \\ 0 & b_2 & b_1 \end{bmatrix} f + \begin{bmatrix} a_3 & a_2 & 0 \\ b_2 & b_1 & x \\ 0 & b_2 & 1 \end{bmatrix} g \in < f, g >$$

■

Definition 11 (Triangular form) *For the sake of simplicity, we will define it for $n = 4$. The triangular form of f is defined as*

$$\tilde{f} = \left(\tilde{f}_1, \dots, \tilde{f}_4\right)$$

where

$$\begin{aligned}\tilde{f}_1 &= R_{x_1,1}\left(f_1, f_2\right)\\ \tilde{f}_2 &= R_{x_2,1}\left(f_{12}, f_{13}\right)\\ \tilde{f}_3 &= R_{x_3,1}\left(f_{23}, f_{24}\right)\\ \tilde{f}_4 &= f_{34}\end{aligned}$$

where again

$$\begin{array}{lll} f_{12} = R_{x_1,0}\left(f_1, f_2\right) & & \\ f_{13} = R_{x_1,0}\left(f_1, f_3\right) & f_{23} = R_{x_2,0}\left(f_{12}, f_{13}\right) & \\ f_{14} = R_{x_1,0}\left(f_1, f_4\right) & f_{24} = R_{x_2,0}\left(f_{12}, f_{14}\right) & f_{34} = R_{x_3,0}\left(f_{23}, f_{24}\right) \end{array}$$

Definition 12 (Near-diagonal form) *The near diagonal form of f is defined as $u \in \mathbb{C}\left[x_n\right]$ and $p \in \mathbb{C}\left(t\right)^n$ obtained from back-substitution from $\tilde{f}$ so that*

$$\begin{aligned}u &= \tilde{f}_n\\ x_1 &= p_1\left(x_n\right)\\ &\vdots\\ x_{n-1} &= p_{n-1}\left(x_n\right)\end{aligned}$$

Theorem 13 *Let* $f \in \mathbb{C}[x_1, \ldots, x_n]^n$ *and let* $g \in \mathbb{C}[x_1, \ldots, x_n]$ *be random linear. Let* $u \in \mathbb{C}[t]$ *and* $p \in \mathbb{C}(t)^n$ *be the near-diagonal form of* $(f_1, \ldots, f_n, g - t)$*. Then we have*

$$S \subset \{p(t) : u(t) = 0\}$$

Proof. Fill in....... ∎

2.2 Algorithm

Algorithm 14 (Triangularize)

In: $f \in \mathbb{C}[x_1, \ldots, x_n]^n$

Out: $f_t \in \mathbb{C}[x_1, \ldots, x_n]^n$ *,triangular form of* f

1. n = *number of equations*
2. ft = *list of resultant*
3. $ft[i] = R_{x_1}(f_1, f_2)(x_2, x_3)$
4. $ft[i+1] = R_{x_1}(f_1, f_3)(x_2, x_3)$
5. *Repeat resultant* n *times*
6. *Return* ft

Algorithm 15 (Near-diagonalize)

In: $f \in \mathbb{C}[x_1, \ldots, x_n]^n$

Out: $f_t \in \mathbb{C}[x_1, \ldots, x_n]^n$ *,near-diagonal form of* f

1. $ft = Triangularize(fs, vs)$
2. *Solve for* $g_n(x_n)$
3. *Back substitute* $g_1, g_2, \ldots, g_{n-1}$
4. *Return list of* g

Algorithm 16 (SolveByUniRes)

In: $f \in \mathbb{C}[x_1, \ldots, x_n]^n$, $g \in \mathbb{C}[x_1, \ldots, x_n]$ *random linear*

Out: $S \subset \mathbb{C}^n$*, the set of complex solutions of* f *(possibly with extraneous ones)*

1. $g = NearDiagonalize(f)$
2. t = *solution of* g
3. S = *back substitution of* t
4. *Return* S

3 Solving by Multivariate Resultant (General systems)

3.1 Theory

Definition 17 (Macaulay matrix) *Let* $f \in \mathbb{C}[x_1, \ldots, x_n]^{n+1}$. *Let*

1. $D = (d_1 - 1) + \cdots + (d_{n+1} - 1) + 1$ *where* $d_i = \deg f_i$
2. $T = \{x^e : |e| \leq D\}$, *ordered in the decreasing order in the total degree where* $x_n > x_{n-1} > \cdots > x_1$.
3. $T_1 = \left\{t \in T : x_1^{d_1} | t\right\}$

   $T_2 = \left\{t \in T \backslash T_1 : x_2^{d_2} | t\right\}$

   $\vdots$

   $T_n = \left\{t \in T \backslash T_1 \backslash \cdots \backslash T_{n-1} : x_n^{d_n} | t\right\}$

   $T_{n+1} = T \backslash T_1 \backslash \cdots \backslash T_n$

The Macaulay matrix M *of* f *is defined such that*

$$\begin{bmatrix} T_1/x_1^{d_1} \ f_1 \\ \vdots \\ T_n/x_n^{d_n} \ f_n \\ T_{n+1} \ f_{n+1} \end{bmatrix} = M \begin{bmatrix} T_1 \\ \vdots \\ T_n \\ T_{n+1} \end{bmatrix}$$

Theorem 18 *If* f *has a common solution then* $|M| = 0$.

Proof. Fill in..... ■

Definition 19 (Eigen matrix) *Let* $f \in \mathbb{C}[x_1, \ldots, x_n]^n$ *and* $g \in \mathbb{C}[x_1, \ldots, x_n]$. *Let* M *be the Macaulay matrix of* $(f_1, \ldots, f_n, g)$. *Let* $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ *where* D *is square with* $d_1 \cdots d_n$. *The Eigen-matrix* E *of* g *modulo* f *is defined by* $E = -CA^{-1}B + D$.

Theorem 20 *Let* $g \in \mathbb{C}[x_1, \ldots, x_n]$ *be random linear. Let* E *be the eigen-matrix of of* g *modulo* f. *Let* V *be the set of all the eigenvectors of* E. *Then, with probabilities one, we have*

$$S = \{[v_{m-1}/v_m, \ldots, v_{m-n}/v_m] : v \in V\}$$

Proof. From class we proved that $T_n(\alpha)$ is an eigenvector of E and we obtain the result

$$\begin{bmatrix} \vdots \\ \vdots \\ v_{m-1} \\ v_m \end{bmatrix} \propto \begin{bmatrix} \vdots \\ \vdots \\ \alpha_1 \\ 1 \end{bmatrix}$$

Hence $\alpha_1 = v_{m-1}/v_m, \alpha_2 = v_{m-2}/v_m \ldots \alpha_n = v_{m-n}/v_m$, and we get

$$S = \{v_{m-1}/v_m, \ldots, v_{m-n}/v_m\}$$

■

3.2 Algorithm

Algorithm 21 (Macaulay matrix)

In: $f \in \mathbb{C}[x_1, \ldots, x_n]^{n+1}$

Out: *The Macaulay matrix of f*

1. $D =$ *number of partition*
2. $D = (degf_1 - 1) + \ldots (degf_n - 1) + 1$
3. $\omega = MonToDeg(D)$
4. *Partition* ω *into* $T_1, \ldots, T_n$
5. *Construct matrix* M *from* $T_1, \ldots, T_n$
6. *Return* M

Algorithm 22 (Eigen matrix)

In: $f \in \mathbb{C}[x_1, \ldots, x_n]^n$, $g \in \mathbb{C}[x_1, \ldots, x_n]$

Out: *E, the eigen matrix of g modulo f*

1. $M = MacaulayMatrix(f)$
2. $E = -CA^{-1}B + D$
3. *Return* E

Algorithm 23 (SolveByMultiRes)

In: $f \in \mathbb{C}[x_1, \ldots, x_n]^n$, $g \in \mathbb{C}[x_1, \ldots, x_n]$ *random linear*

Out: *$S \subset \mathbb{C}^n$, the set of complex solutions of f*

1. $E = EigenMatrix(f)$
2. $V =$ *eigenvectors of* E
3. $S = \{v_{m-1}/v_m, \ldots, v_{m-n}/v_m\}$
4. *Return* S