Gröbner Bases and Holonomic Gradient Method — Evaluation of *A*-Hypergeometric Polynomials

Nobuki Takayama (arxiv:1212.6103(Hibi-Nishiyama-T), 1505.02947(Ohara-T))

hgm OpenXM search.

Let $A = (a_{ij})$ be a $d \times n$ matrix ($a_{ij} \in \mathbf{Z}$). We denote by $a_j \in \mathbf{Z}^d$ the j-th column vector of A. We assume that there exists a row i such that $a_{ij} > 0$. For $\beta \in \mathbf{N}_0 A = \mathbf{N}_0 a_1 + \cdots + \mathbf{N}_0 a_n$, the polynomial

$$Z_A(\beta; x) = \sum_{Au=\beta, u \in \mathbf{N}_0^d} \frac{x^u}{u!} = \sum_{Au=\beta, u \in \mathbf{N}_0^d} \frac{\prod x_i^{u_i}}{\prod u_i!} \qquad (1)$$

is called the A-hypergeometric polynomial [6].
$P(U = u) = \frac{x^u}{u!} / Z_A(\beta; x)$ is a probability distributioin on $Au = \beta$ with a parameter vector x.
Goal: Exact numerical evaluation of the polynomial $Z_A(\beta; x)$ and its derivatives.
This problem is fundamental and has a lot of applications, e.g., $E[U_i] = x_i \frac{\partial Z}{\partial x_i} / Z$.

$$Z_A(\beta; x) = \sum_{Au=\beta, u\in\mathbf{N}_0^d} \frac{x^u}{u!}$$

Example(2 × 2 contingency table): $A = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$,

$\beta = (37, 36, 12)^T$. We denote u by $\begin{pmatrix} u_1 & u_2 \\ u_3 & u_4 \end{pmatrix}$. u's satisfying $Au = \beta$ are

$$u = \begin{pmatrix} 11 & 0 \\ 25 & 12 \end{pmatrix}, \ldots, u = \begin{pmatrix} 4 & 7 \\ 32 & 5 \end{pmatrix}, \ldots, u = \begin{pmatrix} 0 & 11 \\ 36 & 1 \end{pmatrix}.$$

$$Z_A(\beta; x) = \frac{x_2^{11} x_3^{36} x_4}{11!36!1!} {}_2F_1(-12, -11, 26; y), y = \frac{x_1 x_4}{x_2 x_3}.$$

The polynomial Z_A satisfies the A-hypergeometric system (Gel'fand-Kapranov-Zelevinsky hypergeometric system): a_i spans $\mathbf{Z}^d$. $c_1, \ldots, c_d$: indeterminates.

$$D[c] = \mathbf{C}[c_1, \ldots, c_d]\langle x_1, \ldots, x_n, \partial_1, \ldots, \partial_n \rangle$$

where $\partial_i x_j = x_j \partial_i + \delta_{ij}$. $H_A[c]$ is the left ideal in $D[c]$ generated by

$$\sum_{j=1}^{n} a_{ij} x_j \partial_j - c_i =: E_i - c_i, \qquad (i = 1, \ldots, d) \tag{2}$$

$$\prod_{i=1}^{n} \partial_i^{u_i} - \prod_{j=1}^{n} \partial_j^{v_j} \tag{3}$$

(u, v runs over all $u, v \in \mathbf{N}_0^n$ satisfying $Au = Av$.)

The ideal generated by (3) is I_A (the affine toric ideal). For $\beta \in \mathbf{N}_0 A$, the left ideal (generated by) $H_A[\beta]$ (in D), which is called the A-hypergeometric system $H_A(\beta)$, annihilates the polynomial $Z_A(\beta; x)$.

Contiguity relation/Recurrence relation

$$\partial_i \bullet Z_A(\beta; x) = Z_A(\beta - a_i; x)$$

(the contiguity relation)
Numerical evaluation of hypergeometric polynomial becomes hard problem when $\dim \mathrm{Ker}\, A$ and the rank of $H_A(\beta)$ increase and β becomes larger.

Example:

$$F_C(a, b, c; y) = \sum_{k \in \mathbf{N}_0^n} \frac{(a)_{|k|}(b)_{|k|}}{\prod k_i! \prod (c_i)_{k_i}} y^k, \quad A = \begin{pmatrix} \mathbf{1} & \mathbf{1} \\ E_{n+1} & -E_{n+1} \end{pmatrix}$$

where $(a)_m = a(a+1)\cdots(a+m-1)$ and $|k| = k_1 + \cdots + k_n$.
$n = 4,\ a = -179 - N,\ b = -139 - N,\ c = (37, 23, 13, 31)$,
$y = (31/64, 357/800, 51/320, 87/160)$

N	Evaluating series	method of Macaulay type matrix
0	6822s (1.89 hour)	61399s (about 17 hours)
100	138640s (1 day and about 14.5 h)	73126s(about 20.3 hours)
200	More than 2 days	84562s (about 23.5 hours)

```
N=200
A=[[1,0,0,1,0,1,0,1,0,1],[0,1,0,1,0,1,0,1,0,1],[0,0,1,-1,0,0,0,0,0,0],[0,0,0,0,1,-1,0,0,0,0],[0,0,0,0,0,0
Beta=[452,412,-37,-23,-13,31]
at ([x1,x2,x3,x4,x5,x6,x7,x8,x9,x10]=[140/411,40/137,25/822,31/411,14/411,17/274,17/822,5/137,10/137,29/82
oohg_native=0, oohg_curl=1
EV([x3])=[48401824047172895382220332055338065321948101264386648720104327220455411642733594253492395373436
86365699839168924385947529623435213755551773022215922104722152504652845614751116627622765024345097422807
30575009219352322931316768516157628620146639946648721346938153566373438419388097474182951426132409623333
34427535082203520313105491672681943516517877832538986600002769954889790599348816719639272827773538373088
/194422284984251555304384242912588859511600655333063789436840056072076800834495255696040312940357668265848
20636859057551023139439540444360178054580858641760937317843818981263740587028035356318196511904938764035
94177251448953319474978174684020870567460600887603173428867153247620070185651601195645159726853837993587
32090627201429825951569856280808639609886906110220425511570638764915578591464428000430220868340937739443
95739320563272060302627219120238104637235693522860634139129980778711915069111]
Time=84562.4
```

N	Evaluating of series	method of Macaulay type matrix
0	6822s (1.89 hour)	61399s (about 17 hours)
100	138640s (1 day and about 14.5 h)	73126s(about 20.3 hours)
200	More than 2 days	84562s (about 23.5 hours)

Intel Xeon E5-4650 (2.7GHz) with 256G memory, the computer algebra system Risa/Asir (20140528).

Method: the holonomic gradient method (HGM) consisting of 3 steps.

hgm OpenXM search.

The method of Macaulay type matrix is a variation of the HGM.

hgm OpenXM search.

Step 1. (Find a holonomic system for an integral or a sum.) Derive a Pfaffian system for the holonomic system $H_A[c]$.

$$R_n = \mathbf{C}(c, x)\langle \partial_1, \ldots, \partial_n \rangle. \tag{4}$$

$R_n H_A[c]$ is a zero dimensional ideal in R_n. G : a Gröbner basis of the ideal. $\{s_1, \ldots, s_r\}$: the set of the standard monomials for G. r : the rank of $H_A[c]$ (See [1] as to the rank of $H_A(\beta)$.)
$S = (s_1, \ldots, s_r)^T$: the column vector of the standard monomials.
The matrix P_i satisfying $\partial_i S \equiv P_i(c, x) S \bmod R_n H_A[c]$ can be obtained by the normal form computation of $\partial_i s_j$ by G. Y : a column vector of r unknown functions r.

$$\partial_i \bullet Y = \frac{\partial Y}{\partial x_i} = P_i Y \tag{5}$$

is called the Pfaffian system. $Y(\beta; x) = (s_1 \bullet Z_A, \ldots, s_r \bullet Z_A)^T$ satisfies (5). From the contiguity relation, we have

$$Y(\beta - a_i; x) = P_i(\beta, x) Y(\beta; x) \tag{6}$$

$$Y(\beta - a_i; x) = P_i(\beta, x) Y(\beta; x)$$

Example (2×2 contingency table) :

$$Y(\beta - a_1) = \frac{1}{x_1} \begin{pmatrix} -B & -x_2 \\ \frac{\beta_3 B x_3}{D} & \frac{\beta_2 x_1 x_4 + \beta_3 x_2 x_3}{D} \end{pmatrix} Y(\beta; x), \quad D = x_1 x_4 - x_2 x_3, B = \beta_1 - \beta_2 - \beta_3$$

Step 2. Evaluate $Y(\beta'; x)$ for a small β' by evaluating the hypergeometric polynonomial and its derivatives.

Step 3. Extend the value of Y by applying the following relation repeatedly

$$Y(\beta''; x) = P_i(\beta'', x)^{-1} Y(\beta'' - a_i; x) \quad (7)$$

along a suitable path from β' to β.

Difficulties of the “generic” or “general” HGM above are

1. Computation of the Gröbner basis of $R_n H_A[c]$ and normal forms is hard. For example, the normal form algorithm stops with a memory exhaustion in our example.
2. We need to find a suitable path from β' to β. The denominator of the matrix P_i might be zero for a β'' on the path and the matrix P_i might not have the inverse for a β'' on the path.

Solution to the difficulty 2.
A basis S of $R_n/(R_n H_A[c])$ as a vector space over $\mathbf{C}(c, x)$ is called a *good basis* for $\mathbf{N}_0 A$ when the following two conditions are satisfied.

1. The singularity polynomial for the Pfaffian system does not contain the variables c_i's.
2. S is a basis of $R'/R' H_A(\beta)$, $R' = \mathbf{C}(x)\langle \partial_1, \ldots, \partial_n \rangle$ for all $\beta \in \mathbf{N}_0 A$.

Theorem (Ohara-T [4])

When A is normal and S is a good basis, then the matrix $P_i(\beta + a_i; x)$ has the inverse for β satisfying $\beta - \sum u_i a_i \in \mathbf{N}_0 A$ for all $\partial^u \in S$ and for any x out of a measure zero set.

Solution to the difficulty 1.

1-a: A basis S is determined by the following theorem.

Theorem (Hibi-Nishiyama-T [2])

Let $w \in \mathbf{Z}^n$ be a generic weight vector for the affine toric ideal I_A such that $\deg \mathrm{in}_w(I_A) = \deg I_A$. Let $u_1, \ldots, u_r$ be a monomial basis of $R_n/(R_n J)$ where the left ideal J is generated by $\mathrm{in}_w(I_A)$ and $E_i - c_i$, $i = 1, \ldots, d$ in R_n. Then, $\{u_i\}$ is a basis of $R_n/(R_n H_A[c])$.

1-b: Use of Macaulay type matrix [3], which is a generalization of the Sylvester matrix. We construct the Macaulay type matrix F_T from the generators of the ideal $R_n H_A[c]$ and the basis S up to the degree T. We simplify the matrix F_T by a Gröbner basis of the toric ideal I_A and construct a simplifed matrix F'_T. Note that we no longer need non-commutative multiplication to construct the Pfaffian system $P_i(c; x)$ from F'_T. We specialize the parameter c to the vector of natural numbers β'' and the variable x to the evaluation point ξ and construct the Pfaffian system $P_i(\beta''; \xi)$ by computing the echelon form of F'_T.

Theorem (Ohara-T [4])

1. *When T is sufficiently large, the reduced row echelon form of $F' = F'_T$ contains the Pfaffian operator $\partial_i s - \sum_{t \in S} c_t t$ if $\partial_i s$ is irreducible by the Gröbner basis of I_A.*
2. *Let f be a solution of $H_A(c)$. The numerical value $(\partial_i s) \bullet f$ at a generic point $x = \xi \in \mathbf{Q}^n$, $c = \beta'' \in \mathbf{Q}^d$, can be obtained from the numerical values of $s \bullet f$, $s \in S$ at a point $x = \xi, c = \beta''$ by computing the reduced row echelon form of the numerical matrix $F'|_{x=\xi, c=\beta''}$.*

The Macaulay type algorithm is implemented in the package `ot_hgm_ahg.rr` for Risa/Asir.

Note: For the two way contingency tables, a construction of Pfaffian systems by utilyzing twisted cohomology groups is studied by Y.Goto and K.Matsumoto. See [5] and its references for the case of $2 \times m$ contingency tables.

[1] L.Matusevich, E.Miller, U.Walther, Homological Methods for Hypergeometric Families, Journal of American Mathematical Society 18 (2005), 919–941.

[2] T.Hibi, K.Nishiyama, N.Takayama, Pfaffian Systems of *A*-Hypergeometric Systems I, Bases of Twisted Cohomology Groups. arxiv:1212.6103

[3] F.S.Macaulay, On Some Formulae in Elimination, Proceedings of the London Mathematical Society 33 (1902), 3–27.

[4] K.Ohara, N.Takayama, Pfaffian Systems of A-Hypergeometric Systems II — Holonomic Gradient Method, arxiv:1505.02947.

[5] M.Ogawa, Algebraic Statistical Methods for Conditional Inference of Discrete Statistical Models, PhD. Thesis, University of Tokyo, 2015.

[6] M.Saito, B.Sturmfels, N.Takayama, Hypergeometric polynomials and Integer Programming, Compositio Mathematica, 115, (1999) 185–204.

Example: Consider the A-hypergeometric ideal generated by

$$x_1\partial_1 + x_2\partial_2 + x_3\partial_3 + x_4\partial_4 - c_1, x_2\partial_2 + x_4\partial_4 - c_2, x_3\partial_3 + x_4\partial_4 - c_3,$$
$$\underline{\partial_2\partial_3} - \partial_1\partial_4.$$

For the graded reverse lexicographic order such that $\partial_1 > \partial_2 > \partial_3 > \partial_4$, the basis by the algorithm [2] is $S = \{1, \partial_4\}$. Put the degree $T = 1$. We multiply the monomials in $N_T = \{1, \partial_1, \partial_2, \partial_3, \partial_4\}$ to the generators. The table below is the result of the multiplication where the index $i_1 i_2 \cdots i_m$ in the top row stands for the monomial $\prod_{k=1}^m \partial_{i_k}$ and the index 0 denotes the monomial 1.

M'					M_r	M'							S	
11	12	13	14	22	23	24	33	34	44	1	2	3	4	0
										x_1	x_2	x_3	x_4	$-c_1$
x_1	x_2	x_3	x_4							$1-c_1$				
	x_1			x_2	x_3	x_4					$1-c_1$			
		x_1			x_2		x_3	x_4				$1-c_1$		
			x_1			x_2		x_3	x_4				$1-c_1$	
											x_2		x_4	$-c_2$
	x_2		x_4							$-c_2$				
				x_2		x_4					$1-c_2$			
					x_2			x_4				$-c_2$		
						x_2			x_4				$1-c_2$	
												x_3	x_4	$-c_3$
		x_3	x_4							$-c_3$				
					x_3	x_4					$-c_3$			
							x_3	x_4				$1-c_3$		
								x_3	x_4				$1-c_3$	
			-1		1									

and

$M_i \setminus M'$	M_r	$M_i \setminus M'$			M_r		
114	123	124	134	144	223	233	234
-1	1						
		-1			1		
			-1			1	
				-1			1

We put

$$
\begin{aligned}
M_t &= \{23, 114, 123, 124, 134, 144, 223, 233, 234\} \\
M' &= \{1, 2, 3, 11, 12, 13, 14, 22, 24, 33, 34, 44\}.
\end{aligned}
$$

Note that $M_t = M_r \cup (M_i \setminus M')$ in the proof. The join of the two tables is the matrix F and $M = M_t \cup M'$.

Let us use the order of indices such that $M_t > M' > S$. Apply the Gaussian elimination with this order to the first table. We eliminate elements of the column standing for the index 23 by the last row and then remove the last row. Then, we obtain the following matrix which agrees with the matrix F'_T, $(T = 1)$ in the algorithm.

11	12	13	14	22	23	24	33	34	44	1	2	3	4	0
										x_1	x_2	x_3	x_4	$-c_1$
x_1	x_2	x_3	x_4							$1-c_1$				
	x_1		$-x_3$	x_2		x_4					$1-c_1$			
		x_1	$-x_2$				x_3	x_4				$1-c_1$		
			x_1			x_2		x_3	x_4				$1-c_1$	
											x_2		x_4	$-c_2$
	x_2		x_4							$-c_2$				
				x_2		x_4					$1-c_2$			
			$-x_2$					x_4				$-c_2$		
						x_2			x_4				$1-c_2$	
												x_3	x_4	$-c_3$
		x_3	x_4							$-c_3$				
			$-x_3$			x_4					$-c_3$			
							x_3	x_4				$1-c_3$		
								x_3	x_4				$1-c_3$	

We can see, by a calculation, that this matrix can be transformed into the reduced row echelon form whose rank is 12. The reduced row echelon form contains the Pfaffian operators.