Hypergeometric Systems, Weyl Algebra, Creation Operator, and Sequential Conditional Inference of Contingency Tables

Toshio Sakata
Kyushu Institute of Design, Industrial Design
4-9-1 Shiobaru Minami-ku
Fukuoka, Japan
sakata@kyushu-id.ac.jp

Ryuichi Sawae
Okayama University of Science, Applied Mathematics
1-1 Ridai-chyo
Okayama, Japan
sawae@xmath.ous.ac.jp

1. Creation operator

It is very common to use the conditional inference about contingency tables. Here we restrict ourselves to two-way $K \times L$ contingency tables. With a fixed row vector $r = (r_1, ...r_K)$ and a column vector $c = (c_1, ..., c_L)$ let $\Omega(r, c)$ be the set of all contingency table $n = (n_{ij})$ with row vector r and column vector c. Also let define the function $\Phi(x|r, c)$ in the variables $x = (x_{ij})$ as follows,

$$\Phi(x|r,c) = \sum_{n \in \Omega(r,c)} \prod_{i=1}^{K} \prod_{i=1}^{L} \frac{x_{ij}^{n_{ij}}}{n_{ij}}. \tag{1}$$

It is easy to check that $\Phi(x|r, c)$ satisfies

$$\frac{\partial}{\partial x_{ij}} \Phi(x|r,c) = \Phi(x|r - e_i, c - e_j), \tag{2}$$

where $e_i(e_j)$ denotes the unit vector with one in the $i(j)$-th position.. This means the differential operator $\partial_{ij} = \frac{\partial}{\partial x_{ij}}$ reduces one from the i-th row sum and j th column sum. The inverse operator called creation operator is defined by

$$C_{ij} : x_{ij} + \sum_{p=1}^{K} \sum_{q=1}^{L} x_{pj} x_{iq} \frac{\partial}{\partial x_{pq}}. \tag{3}$$

Then it holds that

$$C_{ij} \Phi(x|r,c) = (1 + r_i)(1 + c_j) \Phi(x|r + e_i, c + e_j) \tag{4}$$

This equation implies that, if we apply the creation operators C_{ij}, we can construct the set of all contingency tables with given row sum vector and column sum vector from the set of contingency tables with smaller row sum vector and column sum vector. In fact, starting from the initial function $\Phi(x|0,0) = 1$, we can construct any function $\Phi(x|r,c)$, and thus we can obtain the set $\Omega(r,c)$ for any r and c. It should be noted that the function $\Phi(x|r,c)$ is the polynomial solution of a hypergeometric system of partial differential equations. The theory of creation operator is developed in the context. Note that the creation operators are calculated by using the Groebner basis of Weyl algebra of differential operators. Also note that the theory was applied to the problem of integer programming in Saito et al(1999).

2. Sequential conditional test

Let's assume that we have a contingency table with a row sum vector r and a column sum vector c. Then we can calculate the function $\Phi(x|r,c)$. If we get an added sample at the (i,j) cell then apply the creation operator C_{ij} to the function $\Phi(x|r,c)$, and we get the function $\Phi(x|r+e_i, c+e_j)$, which is the generating function corresponding to the set of contingency tables $\Omega(r+e_i, c+e_j)$, which is the new set of contingency tables needed for the conditional test at the next stage of the sequential test. That is, we can obtain the set of contingency tables with given marginals at each step by applying a creation operator depending on each added sample. Thus we can perform the sequential conditional test. This application of creation operator to the sequential test of contingency tables first appeared in Sakata and Sawae(2000) for discrimination of two 4-nomial distributions. In this paper we study the performance of the sequential test for testing the hypothesis of equality of binomial distributions against the simple order hypothesis.

REFERENCES

M. Saito, B. Strumfels and N. Takayama(1999), Hypergeometric Polynomials and Integer Programming, *Compositio Mathematica*, **115**, 185-204.

T. Sakata and R.Sawae(2000). A new development of the conditional inference of contingency tables *Proc. The 7-th Japan-China Symposium on Statistics*,271-274,

T. Sasaki(1991), Contiguity relations of Aomoto-Gelfand's hypergeometric functionsand applications to Apell's system F_3 and Goursat's system ${}_3F_2$, *SIAM J.of Math. Anal.*, **22**(1991), 821-846.

RESUME

On propose un test conditionnel seuentiel sur des tables de contingence et la performance du test est etudee par la simulation pour le test d'egalit de plusieurs distributions binomiale.