$E(k,n)$ の contiguity relation による乱数生成

高山信毅 (神戸大学)

A-超幾何系の今年のニュース.

1. D-加群: Dual についての予想[*] の証明 (Avi Steiner).
2. 解析: homology cycle とホモロジー交点数 (松原).
3. 統計: “間野の direct sampler” と contiguity relation.

参考文献等 $\boxed{\text{hgm OpenXM}}$ search.

```
http://www.math.kobe-u.
ac.jp/HOME/taka/2018/
nt-tanbara-2018.pdf
```

[*]文献では Takayama 予想

$r_1 \times r_2$ 分割表とその上の分布

1. 2 元分割表とは非負整数を成分とする $r_1 \times r_2$ 行列 u_{ir_1+j} [†].
2. 定数 $p_{ir_1+j} \geq 0$ を (i,j) 成分の確率と呼ぶ.
3. 行和, 列和ベクトルが一定値 β の $r_1 \times r_2$ 分割表に, 確率分布

$$\frac{p^u}{u!}\frac{1}{Z(\beta;p)}, \quad Z(\beta;p) = \sum_{u\text{ の行和列和は }\beta} \frac{p^u}{u!}$$

を考える.

Poisson 分布 $\frac{p^u}{u!}\exp(-p\cdot\mathbf{1})$ に従う random variable U の $AU = \beta$ となる条件付き分布:

$$P(U = u \mid AU = \beta) = \frac{p^u}{u!}\frac{1}{Z(\beta;p)}$$

[†] index のスタートは 0. 1 をスタートとする場合もある. 文脈で適宜判断

例

インフルエンザ脳症で使用した解熱剤と生存率

	アセトアミノフェン	ジクロフェナクナトリウム	計
死亡	4	7	11
生存	32	5	37
計	36	12	48

幾何数理工学と推測数理工学の試験成績

幾何 \ 推測	5	4	3	2	1-	計
5	2	1	1	0	0	4
4	8	3	3	0	0	14
3	0	2	1	1	1	5
2	0	0	0	1	1	2
1-	0	0	0	0	1	1
計	10	6	5	2	3	26

$\boxed{\text{乱数生成}}$ 問題: p と β を与えたときに分布

$$\frac{p^u}{u!}\frac{1}{Z(\beta;p)}$$

に従う乱数 u を生成せよ.

答え 1. Diaconis-Sturmfels 1998. I_A : A[‡] できまる affine toric ideal. この生成元を用いて Markov Chain Monte Carlo 法 (MCMC) を行う[§].

答え 2. 間野 2017, direct sampler. A 超幾何の contiguity を用いて direct sampling する[¶].

[‡] あとで

[§] MCMC 弱点: 品質のよい乱数を生成するには burn-in や thinning が必要.

[¶] 弱点: 遅い. contiguity の導出は toric ideal の generator を求めることよりはるかにむづかしい.

$\boxed{A \text{ による定式化}}$ 非負整数成分の $d \times n$ 行列 $A = (a_{ij})$. A の第 i 列ベクトル a_i を $\mathbf{Z}^d$ の元とみたとき，この列ベクトル達は $\mathbf{Z}^d$ を生成していると仮定する．$a_{1j} = 1$ と仮定.

$\beta \in \mathbf{N}_0 A = \mathbf{N}_0 a_1 + \cdots + \mathbf{N}_0 a_n$ に対して多項式

$$Z_A(\beta; p) = \sum_{Au=\beta, u \in \mathbf{N}_0^n} \frac{p^u}{u!} \tag{1}$$

を A-超幾何多項式とよぶ. $u! = \prod_{i=1}^n u_i!$, $p^u = \prod_{i=1}^n p_i^{u_i}$.

$$\frac{p^u}{u!} \frac{1}{Z_A(\beta; p)}$$

が A-超幾何分布.

問題: p と β を与えたときに A 超幾何分布に従う乱数 u を生成せよ.

例: $A = (1,1)$, $Z_A(\beta; p) = (p_1 + p_2)^\beta / \beta!$. 二項分布に他ならない.

```
hist(rbinom(1000,size=10,prob=1/2));
```

direct sampler アルゴリズム (間野, 2018)

Input: β, p

Output: c

1. $c := (0, 0, \ldots, 0)$ (count vector)
2. $e_i := \frac{p_i Z(\beta - a_i; p)}{\beta_1 Z(\beta; p)}$, $i = 1, \ldots, n^{\|}$.
3. $[0, 1]$ を $e_1 : e_2 : \ldots : e_n$ に分割.
4. $[0, 1]$ に値を持つ一様乱数 t を一つ生成.
5. t が e_j の領域に入ったら c_j を 1 増やす. $\beta := \beta - a_j$.
6. $\beta \neq 0$ なら goto 2.

終了したとき c は $Ac = \beta^{**}$ を満たす.

例: $A = (1, 1)$ の時にためす.

$^{\|}\beta - a_i \notin \mathbf{N}_0^d$ なら $e_i = 0$. $|\beta| = \beta_1 + \ldots + \beta_d$

**計算途中の β でなく input の β

A-超幾何分布となる理由 1

以下 $x = p$.

Proposition

次の恒等式 (2), (3) が成立する.

$$\sum_{i=1}^{n} x_i \frac{\partial Z(\beta; x)}{\partial x_i} = \beta_1 Z(\beta; x) \tag{2}$$

ここで β_1 は β の第一成分である.

$$\frac{\partial Z(\beta; x)}{\partial x_i} = Z(\beta - a_i; x) \tag{3}$$

ここで a_i は 行列 A の i 列目 (長さ d の縦ベクトルとなる) である.

A-超幾何分布となる理由 2

(3) を (2) へ代入して両辺を $\beta_1 Z(\beta; x)$ で割ると次の式を得る.

$$\sum_{i=1}^{n} \frac{x_i Z(\beta - a_i; x)}{\beta_1 Z(\beta; x)} = 1 \tag{4}$$

和が 1 となるので,

$$\frac{x_i Z(\beta - a_i; x)}{\beta_1 Z(\beta; x)}$$

を, 条件 β の元で index i を選択する確率と定義し,

$$P(I = i|\beta) = \frac{x_i Z(\beta - a_i; x)}{\beta_1 Z(\beta; x)} \tag{5}$$

と記すことにする. ここで I は $\{1, 2, \ldots, n\}$ に値をとる random variable である. もちろん

$$\sum_{i=1}^{n} P(I = i|\beta) = 1$$

となっている.

A-超幾何分布となる理由 3　条件 β の元で, 上記の確率に従い index i_1 を選択し, 次に条件 $\beta - a_{i_1}$ の元で, 上記の確率に従い index i_2 を選択し、これを条件が 0 ベクトルになるまで繰り返す. したがって, index の列

$$i_1, i_2, \ldots, i_\ell, \quad \ell = \beta_1$$

を得る確率は

$$P(I = i_1|\beta)P(I = i_2|\beta - a_{i_1})P(I = i_3|\beta - a_{i_1} - a_{i_2}) \cdots P(I = i_\ell|\beta - a_{i_1} - a_{i_2} - \cdot$$

となる. この式に (5) を代入すると次の式を得る.

$$\frac{x_{i_1} Z(\beta - a_{i_1}; x)}{\beta_1 Z(\beta; x)} \frac{x_{i_2} Z(\beta - a_{i_1} - a_{i_2}; x)}{(\beta_1 - 1) Z(\beta - a_{i_1}; x)} \frac{x_{i_3} Z(\beta - a_{i_1} - a_{i_2} - a_{i_3}; x)}{(\beta_1 - 2) Z(\beta - a_{i_1} - a_{i_2}; x)}$$
$$\cdots \quad \cdots \frac{x_{i_\ell} Z(\beta - a_{i_1} - a_{i_2} - \cdots - a_{i_\ell}; x)}{1 Z(\beta - a_{i_1} - a_{i_2} - \cdots - a_{i_{\ell-1}}; x)}$$

$\boxed{A\text{-超幾何分布となる理由 4}}$ 分母分子がキャンセルされてこれは

$$\frac{1}{\beta_1! Z(\beta; x)} \prod_{j=1}^{\ell} x_{i_j} \tag{6}$$

に等しい.

ここで count vector c を導入しよう. index の列 $i_1, i_2, \ldots, i_\ell$ に対して, c_k を index k が index の列に出現する個数とする. たとえば index 列が,

$$2, 2, 1, 2, 1, 1, 3$$

とすれば,

$$c = (3, 3, 1)$$

となる[††]. この時 $\prod_{j=1}^{\ell} x_{i_j} = x^c$ となる. 同じ count vector c を得る index 列は全部で $\frac{(c_1+\cdots+c_n)!}{c!}$ 個ある. たとえば index 列が,

$$2, 2, 1, 2, 1, 1, 3$$

ならば $1, 1, 1, 2, 2, 2, 3$ や $3, 1, 1, 1, 2, 2, 2, 1$ 等も同じ count vector を与える index 列であり, そのようなものは全部で $\frac{7!}{3!3!1!}$ とおりある.

[††] index 1 が 3 個, index 2 が 3 個, index 3 が 1 個.

$\boxed{A\text{-超幾何分布となる理由 5}}$ さて上の考察より count vector c を direct sampler で得る確率は (6) の $\frac{(c_1+\cdots+c_n)!}{c!}$ 倍となる. $\beta_1 = c_1 + \cdots + c_n$ なので, 結局 count vector c を得る確率は A-超幾何分布.

間野 direct sampler の性質

$Z_A(\beta; p)$ を定義から計算するのは β_1 や A が大きいと計算時間の点で困難が増す.

Theorem

1. A-超幾何系の *Pfaffian* はグレブナー基底で計算可能. これを用いると遷移確率 e_i が漸化式で計算できる *.
2. N 個の乱数を計算するための計算量は $O(r^2\beta_1 N)$ プラス *contiguity* の計算量. ここで r は A の *normalized volume*. また有理数の四則の計算量は $O(1)$ とする.

補足:

1. MCMC の計算量は $O(n'(N + (\text{burn-in の回数})))$ †.
2. A 超幾何では Pfaffian = contiguity. Pfaffian の導出は計算量の多い計算であるが, 後藤松本の $E(k,n)$ の contiguity 関係 $\Rightarrow$ 効率的な分割表の sampler ‡.
3. burn-in が不要なので単純に並列化可能.

$^*A = (1,1)$ で漸化式を実演. 一般の実装は `tk_ds_ahg.rr`

$^\dagger n'$ は I_A の生成元の最大次数.

‡実装は `gtt_ds.rr`

A 超幾何系

A : $d \times n$ 行列. 整数成分. A の列ベクトルは a_i. a_i は $\mathbf{Z}^d$ を生成.
$\beta = (\beta_1, \ldots, \beta_d) \in \mathbf{C}^d$ (parameters).

$$\mathbf{C}\langle x_1, \ldots, x_n, \partial_1, \ldots, \partial_n \rangle, \; x_j x_j = x_j x_i, \partial_i \partial_j = \partial_j \partial_i, \partial_i x_j = x_j \partial_i + \delta_{ij}$$

を D または D_n と書く.

Definition

A-hypergeometric system または GKZ hypergeometric system (GKZ, 1989), $H_A(\beta)$, $M_A(\beta) = D_n / H_A(\beta)$:

$$\begin{aligned} (E_i - \beta_i) \bullet f &= 0, \; E_i - \beta_i = \sum_{j=1}^{n} a_{ij} x_j \partial_j - \beta_i, \; (i = 1, \ldots, d) \\ \Box_u \bullet f &= 0, \; \Box_u = \prod_{\{i \,|\, 1 \leq i \leq n, u_i > 0\}} \partial_i^{u_i} - \prod_{\{j \,|\, 1 \leq j \leq n, u_j < 0\}} \partial_j^{-u_j} \end{aligned}$$

with $u \in \mathbf{Z}^n$ running over all u such that $Au = 0, u \neq 0$.
I_A は $\Box_u$ 達が $\mathbf{C}[\partial_1, \ldots, \partial_n]$ で生成するイデアル.

$\boxed{\text{例}}$

$$A(F_C, 2) = \begin{pmatrix} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$\mathrm{degree}(I_A) = \mathrm{vol}(A)$.

Example

Macaulay2 commands to evaluate the volume (the degree) of $A(0134)$. Here, `o5` is I_A.

```
loadPackage "FourTiTwo"
M=matrix "1,1,1,1; 0,1,3,4"
R=QQ[a..d]
I=toricGroebner(M,R)
   o5 = ideal (b^3  - a^2*c, b*c - a*d, - a*c^2  + b^2*d, c^3  - b*d^2
degree(I)
   o6 = 4
```

contigyity と 例

性質: f が $H_A(\beta)$ の解なら, $\partial_i f$ は $H_A(\beta - a_i)$ の解となる.
f および f の偏微分を basis vector F とした Pfaffian

$$\partial_i F = P_i F$$

を作ると, P_i は contiguity

$$P_i(\beta)F(\beta; x) = F(\beta - a_i; x)$$

を与える. $\Rightarrow$ $\boxed{\text{期待値の比 } e_i \text{ の計算が漸化式で可能}}$
例: `A=[[1,1,1],[0,1,2]]`. Pfaffian は

$$\partial_2 - \begin{pmatrix} \frac{\beta_2}{x_1} & -\frac{2x_3}{x_2} \\ \frac{2\beta_2(\beta_2-1)x_1}{4x_2(x_1x_3-x_2^2)} & \frac{-4(\beta_2-1)x_1x_3+(\beta_2-2\beta_1)x_2^2}{4x_2(x_1x_3-x_2^2)} \end{pmatrix}$$

```
load("tk_ds_ahg.rr")$ C=tk_ds_ahg.build_contiguity_0([[1,1,1],[0,1,2]])
```

例: Z の直接計算は時間がかかる 1

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)

例: Z の直接計算は時間がかかる 2

```
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]=[140/411,40/137,25/822,31/411,14/411,17/274,17/822,5/137,10/137,29/8
oohg_native=0, oohg_curl=1
EV([x3])=[48401824047172895382220332055338065321948101264386648720104327220455411642733594253492395373436
86365699839168924385947529623435213755551773022215922104722152504652845614751116627622765024345097422807
30575009219352322931316768516157628620146639946648721346938153566373438419388097474182951426132409623333
34427535082203520313105491672681943516517877832538986600002769954889790599348816719639272827773538373088
/19442228498425155530438424291258885951160065533306378943684005607207680083449525569604031294035766826584
20636859057551023139439540444360178054580858641760937317843818981263740587028035356318196511904938764035
94177251448953319474978174684020870567460600887603173428867153247620070185651601195645159726853837993587
32090627201429825951569856280808639609886906110220425511570638764915578591464428000430220868340937739443
9573932056327206030262721912023810463723569352286063413912998077871191506911]
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).

ソフトウエア gtt_ds.rr, tk_ds_ahg.rr.

```
[1822] load("gtt_ds.rr");
[2720] gtt_ds.direct_sampler([[4,14,3],[10,6,5]],
                  [[1,9/10,11/10],[1,13/10,99/100],[1,1,1]]);
[ 0 1 3 ]
[ 8 5 1 ]
[ 2 0 1 ]
[2721] gtt_ds.direct_sampler([[4,14,3],[10,6,5]],[[1,9/10,11/10],[1,13/
[ 3 1 0 ]
[ 6 4 4 ]
[ 1 1 1 ]
[2722] gtt_ds.direct_sampler([[4,14,3],[10,6,5]],[[1,9/10,11/10],[1,13/
[ 2 1 1 ]
[ 6 4 4 ]
[ 2 1 0 ]
```

$\boxed{\text{アナロジー}}$

$$\frac{I_A \text{ のグレブナー基底} \quad \Rightarrow \quad \text{MCMC が作れる}^{\S}}{A\text{-超幾何 の contiguity} \quad \Rightarrow \quad \text{間野の direct sampler が作れる}^{\P}}.$$

素手で (理論的考察で) contiguity relation を作れれば, random vector を生成する高速アルゴリズムが作れる.

以上. 以下は玉原でご相談したいこと.

---

§JST CREST 日比チーム編, グレブナー道場, 2011, 共立出版

¶S.Mano, Partitions, Hypergeometric Systems, and Dirichlet Processes in Statistics, JSS Research Series in Statistics (2018), Springer

$\boxed{\text{Dual } D\text{-module とは}}$　右 D-ideal J の生成元を

$$\mathrm{ad} \ : \ a_\alpha(x)\partial^\alpha \mapsto (-\partial)^\alpha a_\alpha(x)$$

として (形式随伴), 右 D-加群 D/J を 左 D-加群 $D/\mathrm{ad}(J)$ に変える. $D/\mathrm{ad}(J)$ も $\mathrm{ad}(D/J)$ と書く.
M: holonomic D-module.

$$\mathbf{D}M = \mathrm{ad}(\mathrm{Ext}^n_D(M, D))$$

ここで, $\mathrm{Ext}^n_D(M, D) = \mathrm{H}^n(\mathrm{Hom}_D(F^\bullet, D))$, $F^\bullet$ は M の左 D-加群としての free resolution. 計算機で構成できる[‖].

[‖]Oaku, Takayama, Minimal free resolutions of homogenized D-modules, 2001

$\boxed{A\text{-hg の dual } D\text{-module}}$ 計算機実験で, "dual がまた A-超幾何になりそう" が観察された.

```
loadPackage "Dmodules"
i8 : I=gkz(matrix {{1,1,1},{0,1,2}},{-2,-3});
i9 : N=Ddual I

o9 = cokernel |
  D_2^2-D_1D_3
  x_2D_2+2x_3D_3-2
  x_1D_1-x_3D_3+1
  x_2D_1D_3+2x_3D_2D_3-D_2 2x_1x_3D_2D_3+x_2x_3D_3^2-x_1D_2 x_2^2x_3D_3
```

$\beta = (1,2)^T$. $(1,2)^T \sim (2,3)^T$? hypergeometric b:

$$b(\beta) = \beta_2(2\beta_1 - \beta_2)$$

$H_A(\beta)$ と ∂_2 の生成するイデアルは $b(\beta)$ で生成される.
$U\partial_2 \sim 3(4-3) \mod H_A(2,3)$ なので
$M_A(1,2) \ni 1 \mapsto \frac{U}{3} \in M_A(2,3)$ で同型 (左 D 加群としての contiguity)

b の計算プログラム (asir)

```
import("names.rr")$
L=[x1*dx1+x2*dx2+x3*dx3-b1,
          x2*dx2+2*x3*dx3-b2,
   dx1*dx3-dx2^2]$
V=[b1,b2,x1,x2,x3,db1,db2,dx1,dx2,dx3]$
M=newmat(10,10,[
 [0,0,0,0,0, 0,0,1,1,1],
 [0,0,0,0,0, 0,0,0,0,-1],
 [0,0,0,0,0, 0,0,0,-1,0],
 [0,0,0,0,0, 0,0,-1,0,0],
 [0,0,1,1,1, 0,0,0,0,0],
 [0,0,0,0,-1,0,0,0,0,0],
 [0,0,0,-1,0,0,0,0,0,0],
 [0,0,-1,0,0,0,0,0,0,0],
 [1,1, 0,0,0,0,0,0,0,0],
 [1,0, 0,0,0,0,0,0,0,0]])$
G=nd_gr(append(L,[dx2]),V,0,M)$
fctr(G[0]);
end$
```

$\boxed{A \text{ 超幾何は affine toric variety の上の } D\text{-加群}}$ 形式 Fourier 変換:

$$\mathrm{FL} \ : \ x_i \to -\partial_i, \partial_i \to x_i$$

$$\sum_{j=1}^{n} -a_{ij}\partial_j x_j - \beta_i, \ (i = 1, \ldots, d)$$

$$\prod_{\{i \mid 1 \le i \le n, u_i > 0\}} x_i^{u_i} - \prod_{\{j \mid 1 \le j \le n, u_j < 0\}} x_j^{-u_j}, \ (Au = 0)$$

2 番目の式は affine toric ideal I_A の定義式.

$$(\mathbf{C}^*)^d \ni t \mapsto (t^{a_1}, \ldots, t^{a_n}) \in \mathbf{C}^n$$

の像の閉包が $V(I_A)$. $(\mathbf{C}^*)^d$ に制限して考えると affine toric variety の上のべき関数のみたす方程式に他ならない:
$t_i\partial_{t_i} = \sum_j (a_j)_i x_i \partial_{x_i}$.

torus 上の D-module

1. 斎藤理論の出発点: affine toric variety $V(I_A)$ 上の微分作用素環は A-超幾何系の contiguity relation!! 全部 (遠くまでパラメータを飛ばすものも含む) の代数とほぼ同じもの ($x_i \Leftrightarrow \partial_i$).
2. $(bfC^*)^d \to \mathbf{C}^n$, $O_\tau \to V(I_A)$, $V(I_A) \to \mathbf{C}^n$ に対応する D-modules の種々の functor の具体的構成.
3. $(\mathbf{C}^*)^d \subset U \subset \mathbf{C}^n$ **: torus 上の rank 1 locally constant sheaf の $\mathbf{C}^n$ への種々の拡張. Mixed Gauss-Manin (Avi Steiner, 2017, 2018).
4. A が normal の時, dual についての予想の解決.

** U は Zariski open, torus action で不変