TORIC GEOMETRY SEMINAR 2010
Combinatorial Commutative Algebra, Optimization and Statistics

TORIC MODELS IN STATISTICS
EXPONENTIAL FAMILIES, MARKOV CHAINS, REVERSIBLE MARKOV CHAINS, DAG MODELS, BORDERS, DIFFERENTIAL ALGEBRA

Giovanni Pistone
Collegio Carlo Alberto

giovanni.pistone@gmail.com

November 13, 2010

Abstract

Monomial statistical models of on finite integer sample space have many applications. The case more commonly discussed is exponential families, but many other examples are known. We discuss algebraic statistical models with a monomial parameterization and linear constrains. We give simple examples of

- Markov chains,
- reversible Markov chains,
- DAG models,

together with examples of

- the border of such models and its relation with manipulation,
- differential equations satisfied by the normalizing constant.

The talk is based on joint work in progress with L. Malagò, E. Riccomagno, M.-P Rogantin, H. Wynn.

A-models: a definition?

- Let be given a nonnegative integer matrix $A \in \mathbb{Z}_{\geq}^{m+1,n}$. The elements are denoted by $A_i(j)$, $i = 0 \dots m, j = 1 \dots n$. We assume the row A_0 to be the constant 1. Each row of A is the logarithm of a monomial term denoted $t^{A(j)} = t_0 t_1^{A_1(1)} \cdots t_m^{A_m(j)}$.
- On a finite sample space $\mathcal{X}$ we consider unnormalized probability densities

$$q(x;t) = t^{A(x)}, \quad x \in \mathcal{X}, t \in \mathbb{R}_{\geq}^{m+1}.$$

For each reference measure μ on $\mathcal{X}$ we define the probability density

$$p(x;t) = \frac{t^{A(x)}}{\sum_{x\in\mathcal{X}} t^{A(x)} \ \mu(x)}, \quad x \in \mathcal{X},$$

for all $t \in \mathbb{R}_{\geq}^{m+1}$ such that q_t is not identically zero.
- The parameter t_0 cancels out, i.e the density is parameterized by $t_1 \dots t_m$ only. The unnormalized density is a projective object.

C-constrained A-model; identification

- In some applications the statistical model is further constrained by a matrix $C \in \mathbb{Z}^{k,n}$.
$$\begin{cases} q(x;t) & = & t^{A(x)}, \\ \sum_{x \in \mathcal{X}} C_i(x) q(x;t) & = & 0, \end{cases}$$
for $\quad x \in \mathcal{X}, t \in \mathbb{R}_{\geq}^{m+1}, i = 1 \dots k$.

- Assume $s, t \in \mathbb{R}_{>}^{m}$ and $p_s = p_t$. Denote by Z the normalizing constant. Then $p_t = p_s$ if, and only if,
$$Z(s) t^{A(x)} = Z(t) s^{A(x)}, \quad x \in \mathcal{X}$$
hence
$$\sum_{i=0}^{m} (\log t_i - \log s_i) A_i(x) = \log Z(t) - \log Z(s), \quad x \in \mathcal{X}.$$
The confounding condition is
$$\delta^T A = 1, \quad \delta_i = (\log t_i - \log s_i)/(\log Z(t) - \log Z(s)),$$
so that $\delta \in e_0 + \ker A^T$.

Toric ideals; closure of the A-model

- The ker of the ring homomorphism

$$k[q(x) \colon x \in \mathcal{X}] \ni q(x) \mapsto t^{A(x)} \in k[t_0, \ldots, t_m]$$

is the toric ideal of A, $\mathrm{I}(A)$. It has a finite basis made of binomials of the form

$$\prod_{x \colon u(x) > 0} q(x)^{u^+(x)} - \prod_{x \colon u(x) < 0} q(x)^{u^-(x)}$$

with $u \in \mathbb{Z}^{\mathcal{X}}$, $Au = 0$.

- As $\sum_{x \in \mathcal{X}} u(x) = 0$, all the binomials are homogeneous polynomials so that all densities p_t in the A-model satisfy the same binomial equation.

Theorem

- *The nonnegative part of the A-variety is the (weak) closure of the A-model.*
- *Let $\mathcal{H}$ be the Hilbert basis of* $\mathrm{Span}\,(A_0, A_1, \ldots) \cap Z_{\geq}^{\mathcal{X}}$. *Let H be the matrix whose rows are the elements of $\mathcal{H}$ of minimal support.*

Example: 3 binary identical RVs, no-3-way-interaction

- $\mathcal{X} = \{+, -\}^3$. Matrix A is

	+++	−++	+−+	−−+	++−	−+−	+−−	−−−
I	1	1	1	1	1	1	1	1
1	0	1	0	1	0	1	0	1
2	0	0	1	1	0	0	1	1
3	0	0	0	0	1	1	1	1
12	0	1	1	0	0	1	1	0
13	0	1	0	1	1	0	1	0
23	0	0	1	1	1	1	0	0

- Constrain matrix C is

	+++	−++	+−+	−−+	++−	−+−	+−−	−−−
1=2	0	1	−1	0	0	1	−1	0
1=3	0	1	0	1	−1	0	−1	0

- The toric ideal $\mathrm{I}(A)$ is generated by $q(+++)q(--+)q(-+-)q(+--) - q(-++)q(+-+)q(++-)q(---)$.
- Matrix H is quadratic

1	0	0	0	0	0	0	1
0	0	0	0	1	0	1	0
0	0	1	0	0	0	1	0
0	0	0	1	0	0	0	1
0	1	0	0	0	0	1	0
0	0	0	0	0	1	0	1
0	0	0	0	0	0	1	1
1	1	0	0	0	0	0	0
1	0	1	0	0	0	0	0
1	0	0	0	1	0	0	0
0	1	0	1	0	0	0	0
0	0	1	1	0	0	0	0
0	0	0	1	1	0	0	0
0	1	0	0	0	1	0	0
0	0	1	0	0	1	0	0
0	0	0	0	1	1	0	0

Toric and Weyl

- Consider the design $\mathcal{D} \subset \mathbb{Z}_+^d$ with reference measure μ. Let $I(\mathcal{D})$ be the ideal of points. Consider the statistical model

$$q(x;t) = \prod_{i=1}^{d} t_i^{x_i}, \quad x \in \mathcal{D}, \quad t_j \geq 0, \quad j = 1, \dots, d,$$

with normalizing constant

$$Z(t) = \sum_{x \in \mathcal{D}} t^x \mu(x)$$

It is the A-model with $A_i(x) = x_i$, $i = 1, \dots, m$.

- In the Weyl algebra $\mathbb{C}\langle t_1 \dots t_d, \partial_1 \dots \partial_d \rangle$ define the operators

$$A(i,x) = t_i \partial_i - x_i = \partial_i t_i - (1 + x_i), \quad i = 1, \dots, d, \quad x \in \mathcal{D},$$

where the second equality follows from the commutation relation $\partial_i t_i = 1 + t_i \partial_i$. For all $x \in \mathcal{D}$ we have

$$A(i,x) \bullet t^x = \partial_i \bullet (t_i t^x) - (1 + x_i) t^x = 0,$$

so that $t_i \partial_i \bullet t^x = x_i t^x$ and, by iteration, $(t_i \partial_i)^\alpha \bullet t^x = x_i^\alpha t^x$, $\alpha \in \mathbb{N}$.

- The operator $(t_i\partial_i)^\alpha$ applied to the polynomial $Z(t) \in \mathbb{C}[t_1, \ldots, t_m]$ gives

$$(t_i\partial_i)^\alpha \bullet Z(t) = \sum_{x\in\mathcal{D}} (t_i\partial_i)^\alpha \bullet t^x = \sum_{x\in\mathcal{D}} x_i^\alpha t^x \quad (\mu(x) = 1).$$

- Note the commutativity

$$(t_i\partial_i)(t_j\partial_j) = (t_j\partial_j)(t_i\partial_i),$$

hence

$$\prod_{i=1}^d (t_i\partial_i)^{\alpha_i} \bullet Z(t) = \sum_{x\in\mathcal{D}} \prod_{i=1}^d (t_i\partial_i)^{\alpha_i} \bullet t^x = \sum_{x\in\mathcal{D}} \left(\prod_{i=1}^d x_i^{\alpha_i} \right) t^x.$$

- By dividing by the normalizing constant we obtain he following expression for the moments:

$$Z(t)^{-1} \prod_{i=1}^d (t_i\partial_i)^{\alpha_i} \bullet Z(t) = Z(t)^{-1} \sum_{x\in\mathcal{D}} \prod_{i=1}^d (t_i\partial_i)^{\alpha_i} \bullet t^x = \mathrm{E}_t \left[X^\alpha \right].$$

From the ring homomorphism $A\colon \begin{cases} \mathbb{C}[x] & \to & \mathbb{C}\langle t_1 \dots t_d, \partial_1 \dots \partial_d \rangle \\ x_i & \mapsto & t_i \partial_i \end{cases}$ we have

$$A(f(x)) \bullet Z(t) = \sum_{x \in \mathcal{D}} f(x) t^x .$$

Theorem

1. *Let x^α, $\alpha \in M$, be a monomial basis for $\mathcal{D}$. Then $Z(t)$ satisfies the following system of $\#M = \#\mathcal{D}$ linear non-homogeneous differential equations:*

$$A(x^\alpha) \bullet Z(t) = \sum_{x \in \mathcal{D}} x^\alpha t^x, \quad \alpha \in M.$$

2. *Let $f_a(x)$ be the (reduced) indicator polynomial of $a \in \mathcal{D}$. Then $Z(t)$ satisfies the following system of $\#\mathcal{D}$ linear non-homogeneous differential equations:*

$$A(f_a(x)) \bullet Z(t) = t^a, \quad a \in \mathcal{D}$$

3. *Let $g(p_a \colon a \in \mathcal{D})$ be a polynomial in the toric ideal of the monomial homomorphism $p_a \mapsto t^a$. Then*

$$g\left(A(f_a(x)) \bullet Z(t) \colon a \in \mathcal{D}\right) = 0$$

Directed Acyclic Graphs DAGs

- A famous model is

$$\begin{array}{ccc} \text{Carcinogenic genotype } U & \longrightarrow & Y \text{ Lung cancer} \\ \downarrow & & \uparrow \\ \text{Smoking } X & \longrightarrow & Z \text{ tar deposit on lungs} \end{array}$$

  where U is unobservable, X, Y, Z are observable.

- The DAG, together with the ordering

$$U \prec X \prec Z \prec Y$$

  encodes the factorization of probability

$$p(u, x, z, y) = p_1(u) p_2(x|u) p_3(z|x) p_4(y|u, z),$$

- which, in turn, is equivalent to the following two statements of conditional independence

$$U \perp\!\!\!\perp Z | X \quad X \perp\!\!\!\perp Y | U, Z$$

- It is a constrained A-model.

Intervention

- Assume we force the population to avoid smoking. The intervention hides the influence of U on Z, producing the new DAG $U \longrightarrow Y \longleftarrow X$. The new factorization is

$$p(u,z,y \| X=0) = \begin{cases} p_1(u)p_3(z|0)p_4(y|u,z) & \text{on } \{X=0\} \\ 0 & \text{on } \{X=1\} \end{cases}$$

- The conditional independence statements are equivalent to

$$U \perp\!\!\!\perp Z | X \begin{cases} p(u,x,z,+)p(u',x,z',+) - p(u,x,z',+)p(u',x,z,+) = 0 \\ u,u' \in \Omega_1, x \in \Omega_2, z,z' \in \Omega_3 \quad u \neq u' \text{ and } z \neq z' \end{cases}$$

$$X \perp\!\!\!\perp Y | U,Z \begin{cases} p(u,x,z,y)p(u,x',z,y') - p(u,x,z,y')p(u,x',z,y') = 0 \\ u \in \Omega_1, x,x' \in \Omega_2, z \in \Omega_3, y,y' \in \Omega_4 \quad x \neq x' \text{ and } y \neq y' \end{cases}$$

- Does the intervention rule derive from the equations?

Binary case

- Before intervention:

$$U \perp\!\!\!\perp Z | X \begin{cases} p(0,0,0,+)p(1,0,1,+) - p(0,0,1,+)p(1,0,0,+) = 0 \\ p(0,1,0,+)p(1,1,1,+) - p(0,1,1,+)p(1,1,0,+) = 0 \end{cases}$$

$$X \perp\!\!\!\perp Y | U, Z \begin{cases} p(0,0,0,0)p(0,1,0,1) - p(0,0,0,1)p(0,1,0,0) = 0 \\ p(0,0,1,0)p(0,1,1,1) - p(0,0,1,1)p(0,1,1,0) = 0 \\ p(1,0,0,0)p(1,1,0,1) - p(1,0,0,1)p(1,1,0,0) = 0 \\ p(1,0,1,0)p(1,1,1,1) - p(1,0,1,1)p(1,1,1,0) = 0 \end{cases}$$

- After intervention:

$$U \perp\!\!\!\perp Z | Y \begin{cases} p(0,0,0,0)p(1,0,1,0) - p(0,0,1,0)p(1,0,0,0) = 0 \\ p(0,0,0,1)p(1,0,1,1) - p(0,0,1,1)p(1,0,0,1) = 0 \end{cases}$$

$$\{X = 1\} \begin{cases} p(u,1,z,y) = 0 \\ \text{for } u, z, y = 0, 1 \end{cases}$$

Markov Chains MCs

- In a Markov chain with state space V, initial probability π_0 and stationary transitions $P_{u\to v}$, $u, v \in V$, the joint distribution up to time T on the sample space Ω_T is

$$P(\omega) = \prod_{v \in V} \pi_0(v)^{(X_0(\omega)=v)} \prod_{a \in \mathcal{A}} P_a^{N_a(\omega)}, \qquad \text{(M)}$$

  where $(V, \mathcal{A})$ is the directed graph defined by $u \to v \in \mathcal{A}$ if, and only if, $P_{u\to v} > 0$.

- A MC is an instance of the A model with $m = \#V + \#\mathcal{A}$, $n = \#\Omega_T$ and rows

$$A_0(\omega) = 1, A_v(\omega) = (X_0(\omega) = 1), A_a(\omega) = N_a(\omega)$$

  i.e the unnormalized density is

$$q(\omega; t) = t_0 \prod_{v \in V} t_v^{(X_0(\omega)=v)} \prod_{a \in \mathcal{A}} t_a^{N_a(\omega)} \qquad \text{(A)}$$

- The (MC) model is derived from the (A) model by adding the constrains

$$\sum_{v \in V} t_v = \sum_{v\colon\, u \to v \in \mathcal{A}} q_{u\to v}, \quad u \in V.$$

A-model of a MC

- The unconstrained A-model of the MC is a Markov proces with non-stationary transition probabilities.
- The unconstrained model is described probabilistically as follows. Define $a(v) = \sum_{v \to w \in \mathcal{A}} t_{v \to w}$; hence $P_{u \to v} = t_{v \to w}/a(v)$ is a transition probability. Also $\nu(v) = a(v)/\sum_v a(v)$ is a probability. Consider the change of parameters

$$b\pi(v) = t_v, a\nu(v)P_{v \to w} = t_{v \to w},$$

  to get

$$\begin{aligned} q(\omega;) &= t_0 \prod_{v \in V} (b\pi(v))^{(X_0(\omega)=v)} \prod_{v \to w \in \mathcal{A}} (aa(v)P_{v \to w})^{N_{v \to w}(\omega)} \\ &= t_0 b a^N \prod_{v \in V} \pi(v)^{(X_0(\omega)=v)} \prod_{v \in V} \nu(v)^{N_{v+}} \prod_{v \to w \in \mathcal{A}} P_{v \to w}^{N_{v \to w}(\omega)} \end{aligned}$$

- It is a change in reference measure.

Example: binary state space $V = \{+1, -1\}$

- For $e_1, e_2 = \pm 1$ we have

$$N_{e_1 \to e_2} = \frac{1}{4} \sum_{t=1}^{T} (1 + e_1 X_{t-1})(1 + e_2 X_t)$$
$$= \frac{T}{4} + \frac{e_1}{4} X_0 + \frac{e_2}{4} X_T + \frac{e_1 + e_2}{4} \sum_{t=1}^{T-1} X_t + \frac{e_1 e_2}{4} \sum_{t=1}^{T} X_{t-1} X_t$$

- The orthogonal space to $(X_0 = e_i)$ and to all the transition's counts is orthogonally generated by
  1. all monomial terms X^α, $\alpha \in \{0, 1\}$, $\sum \alpha \geq 3$, i.e. the interactions of order at least 3;
  2. all terms $X_s X_t$, $s + 1 < t$, i.e. all binary non consecutive interactions;
  3. all differences $X_t - X_{t-1}$, $t = 1, \ldots, T$, i.e the standard basis of contrasts;
  4. the final value X_T.

Detailed balance

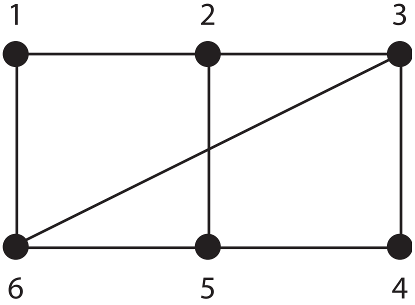
- Consider a simple graph $(V, \mathcal{A})$.
- A transition matrix $P_{v \to w}$, $v, w \in V$, satisfies the detailed balance conditions if $\kappa(v) > 0$, $v \in V$, and

$$\kappa(v) P_{v \to w} = \kappa(w) P_{w \to v}, \quad v \to w \in \mathcal{A}.$$

- It follows that $\pi(v) \propto \kappa(v)$ is an invariant probability and the Markov chain X_n, $n = 0, 1, \ldots,$ has reversible two-step joint distribution

$$\mathrm{P}\left(X_n = v, X_{n+1} = w\right) = \mathrm{P}\left(X_n = w, X_{n+1} = v\right), \quad v, w \in V, n \geq 0.$$

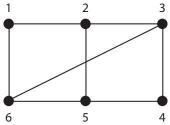
Example: 6 vertexes, 8 edges



$\Gamma =$

	{1, 2}	{2, 3}	{1, 6}	{2, 5}	{3, 4}	{5, 6}	{4, 5}	{3, 6}
1	1	0	1	0	0	0	0	0
2	1	1	0	1	0	0	0	0
3	0	1	0	0	1	0	0	1
4	0	0	0	0	1	0	1	0
5	0	0	0	1	0	1	1	0
6	0	0	1	0	0	1	0	1

CoCoA elimination




```
Use S::=Q[t,k[1..6],p[1..6,1..6]];
Set Indentation;
NI:=6; M:=[];
Define Lista(L,NI);
   For I:=1 To NI Do
       For J:=1 To I-1 Do
           Append(L,k[I]p[I,J]-k[J]p[J,I]); EndFor;
   EndFor; Return L; EndDefine;
N:=Lista(M,NI);
LL:=t*Product([k[I]|I In 1..NI])-1; Append(N,LL);
P0:=[p[1,3],p[1,4],p[1,5],p[2,4],p[2,6], p[3,1],p[3,5],
p[4,1],p[4,2],p[4,6],p[5,1],p[5,3],p[6,2],p[6,4]];
N:=Concat(N,P0);
E:=Elim(k,Ideal(N)); GB:=ReducedGBasis(E);  GB;
```

CoCoA output

```
GB;
[
 p[1,3],  p[1,4],  p[1,5],  p[2,4],  p[2,6],  p[3,1],  p[3,5],
 p[4,1],  p[4,2],  p[4,6],  p[5,1],  p[5,3],  p[6,2],  p[6,4],

  p[2,3]p[3,4]p[4,5]p[5,2] - p[2,5]p[3,2]p[4,3]p[5,4],
  p[1,2]p[2,3]p[3,6]p[6,1] - p[1,6]p[2,1]p[3,2]p[6,3],
  p[1,2]p[2,5]p[5,6]p[6,1] - p[1,6]p[2,1]p[5,2]p[6,5],
  p[2,5]p[3,2]p[5,6]p[6,3] - p[2,3]p[3,6]p[5,2]p[6,5],
  p[3,4]p[4,5]p[5,6]p[6,3] - p[3,6]p[4,3]p[5,4]p[6,5],
  p[1,2]p[2,5]p[3,6]p[4,3]p[5,4]p[6,1] -
        p[1,6]p[2,1]p[3,4]p[4,5]p[5,2]p[6,3],
  p[1,2]p[2,3]p[3,4]p[4,5]p[5,6]p[6,1] -
        p[1,6]p[2,1]p[3,2]p[4,3]p[5,4]p[6,5]]
```

Reversibility on trajectories

Let $\omega = v_0 \cdots v_n$ be a trajectory (path) in the connected graph $\mathcal{G} = (V, \mathcal{E})$ and let $r\omega = v_n \cdots v_0$ be the reversed trajectory.

Proposition

If the detailed balance holds, the the reversibility condition

$$\mathrm{P}(\omega) = \mathrm{P}(r\omega)$$

holds for each trajectory ω.

Proof.

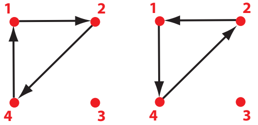
Write the detailed balance along the trajectory,

$$\begin{aligned}
\pi(v_0)P_{v_0 \to v_1} &= \pi(v_1)P_{v_1 \to v_0}, \\
\pi(v_1)P_{v_1 \to v_2} &= \pi(v_2)P_{v_2 \to v_1}, \\
&\vdots \\
\pi(v_{n-1})P_{v_{n-1} \to v_n} &= \pi(v_n)p_{v_n \to v_{n-1}},
\end{aligned}$$

and clear $\pi(v_1) \cdots \pi(v_{n-1})$ in both sides of the product. □

Kolmogorov's condition

We denote by ω a closed trajectory, that is a trajectory on the graph such that the last state coincides with the first one, $\omega = v_0 v_1 \dots v_n v_0$, and by $r\omega$ the reversed trajectory $r\omega = v_0 v_n \dots v_1 v_0$



Theorem (Kolmogorov)

Let the Markov chain $(X_n)_{n \in \mathbb{N}}$ have a transition supported by the connected graph $\mathcal{G}$.

- *If the process is reversible, for all closed trajectory*

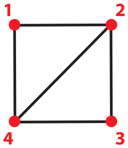
$$\boxed{P_{v_0 \to v_1} \cdots P_{v_n \to v_0} = P_{v_0 \to v_n} \cdots P_{v_1 \to v_0}}$$

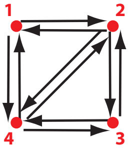
- *If the equality is true for all closed trajectory, then the process is reversible.*

- The Kolmogorov's condition does not involve the π.
- Detailed balance, reversibility, Kolmogorov's condition are algebraic in nature and define binomial ideals.

Transition graph

- From $\mathcal{G} = (V, \mathcal{E})$ an (undirected simple) graph, split each edge into two opposite arcs to get a connected directed graph (without loops) $\mathcal{O} = (V, \mathcal{A})$. The arc going from vertex v to vertex w is $(v \to w)$. The reversed arc is $r(v \to w) = (w \to v)$.

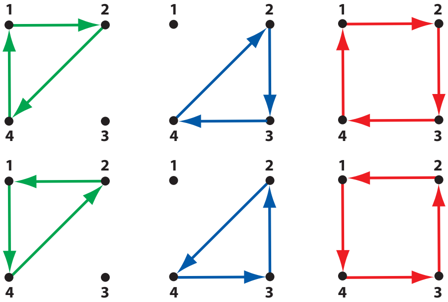




- A path or trajectory is a sequence of vertices $\omega = v_0 v_1 \cdots v_n$ with $(v_{k-1} \to v_k) \in \mathcal{A}$, $k = 1, \ldots, n$. The reversed path is $r\omega = v_n v_{n-1} \cdots v_0$. Equivalently, a path is a sequence of inter-connected arcs $\omega = a_1 \ldots a_n$, $a_k = (v_{k-1} \to v_k)$, and $r\omega = r(a_n) \ldots r(a_1)$.

Circuits, cycles

- A closed path $\omega = v_0 v_1 \cdots v_{n-1} v_0$ is any path going from an initial v_0 back to v_0; $r\omega = v_0 v_{n-1} \cdots v_1 v_0$ is the reversed closed path. If we do not distinguish any initial vertex, the equivalence class of closed paths is called a circuit.
- A closed path is elementary if it has no proper closed sub-path, i.e. if does not meet twice the same vertex except the initial one v_0. The circuit of an elementary closed path is a cycle.



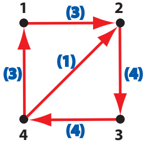
Kolmogorov's ideal

- With indeterminates $P = [P_{v \to w}]$, $(v \to w) \in \mathcal{A}$, form the ring $k[P_{v \to w} : (v \to w) \in \mathcal{A}]$. For a trajectory ω, define the monomial term

$$\omega = a_1 \cdots a_n \mapsto P^{\omega} = \prod_{k=1}^{n} P_{a_k} = \prod_{a \in \mathcal{A}} P_a^{N_a(\omega)},$$

with $N_a(\omega)$ the number of traversals of the arc a by the trajectory.

2→3 3→4 4→1 1→2
2→3 3→4 4→2
2→3 3→4 4→1 1→2
2→3 3→4 4→1 1→2



$P_{1\to2}^{3} P_{2\to3}^{4} P_{3\to4}^{4} P_{4\to1}^{3} P_{4\to2}$

Definition (K-ideal)

The Kolmogorov's ideal or K-ideal of the graph $\mathcal{G}$ is the ideal generated by the binomials $P^{\omega} - P^{r\omega}$, where ω is any circuit.

Bases of the K-ideal

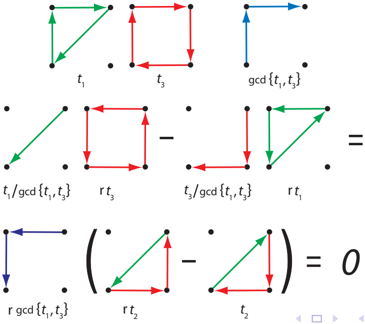
Finite basis of the K-ideal

The K-ideal is generated by the set of binomials $P^{\omega} - P^{r\omega}$, where ω is cycle.

Universal G-basis

The binomials $P^{\omega} - P^{r\omega}$, where ω is any cycle, form a reduced universal Gröbner basis of the K-ideal.

Six cycles: $\omega_1 = 1 \to 2\ 2 \to 4\ 4 \to 1$ (green), $\omega_2 = 2 \to 3\ 3 \to 4\ 4 \to 2$, $\omega_3 = 1 \to 2\ 2 \to 3\ 3 \to 2\ 4 \to 1$ (red), $\omega_4 = r\omega_1$, $\omega_5 = r\omega_2$,$\omega_6 = r\omega_3$.



Cycle space of $\mathcal{O}$

- For each cycle ω define cycle vector

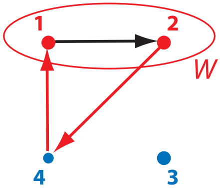
$$z_a(\omega) = \begin{cases} +1 & \text{if } a \text{ is an arc of } \omega, \\ -1 & \text{if } r(a) \text{ is an arc of } \omega, \\ 0 & \text{otherwise.} \end{cases} \qquad a \in \mathcal{A}.$$

- The binomial $P^\omega - P^{r\omega}$ is written as $P^{z^+(\omega)} - P^{z^-(\omega)}$.
- The definition of z can be is extended to any circuit $\bar{\omega}$ by $z_a(\bar{\omega}) = N_a(\omega) - N_a(r\omega)$.
- There exists a sequence of cycles such that $z(\bar{\omega}) = z(\omega_1) + \cdots + z(\omega_l)$.
- We can find nonnegative integers $\lambda(\omega)$ such that $z(\bar{\omega}) = \sum_{\omega \in \mathcal{C}} \lambda(\omega) z(\omega)$, i.e. it belongs to the integer lattice generated by the cycle vectors.
- $Z(\mathcal{O})$ is the cycle space, i.e. the vector space generated in $k^{\mathcal{A}}$ by the cycle vectors.

Cocycle space of $\mathcal{O}$

- For each subset W of V, define cocycle vector

$$u_a(W) = \begin{cases} +1 & \text{if } a \text{ exits from } W, \\ -1 & \text{if } a \text{ enters into } W, \\ 0 & \text{otherwise.} \end{cases} \qquad a \in \mathcal{A}.$$



- The generated subspace of $k^{\mathcal{A}}$ is the cocycle space $U(\mathcal{O})$
- The cycle space and the cocycle space orthogonally split the vector space $\{y \in k^{\mathcal{A}} : y_a = -y_{r(a)}, a \in \mathcal{A}\}$.
- Note that for each cycle vector $z(\omega)$, cocycle vector $u(W)$, $z_a(\omega)u_a(W) = z_{r(a)}(\omega)u_{r(a)}(W)$, $a \in \mathcal{A}$, hence

$$z(\omega) \cdot u(W) = 2\sum_{a\in\omega} u_a(W) = 2\left[\sum_{a\in\omega, u_a(W)=+1} 1 - \sum_{a\in\omega, u_a(W)=-1} 1\right] = 0.$$

Toric ideals

- Let U be the matrix whose rows are the cocycle vectors $u(W)$, $W \subset V$. We call the matrix $U = [u_a(W)]_{W \subset V, a \in \mathcal{A}}$ the cocycle matrix.
- Consider the ring $k[P_a \colon a \in \mathcal{A}]$ and the Laurent ring $k(t_W \colon W \subset V)$, together with their homomorphism h defined by

$$h \colon P_a \longmapsto \prod_{W \subset V} t_W^{u_a(W)} = t^{u_a}.$$

- The kernel $I(U)$ of h is the toric ideal of U. It is a prime ideal and the binomials $P^{z^+} - P^{z^-}$, $z \in \mathbb{Z}^{\mathcal{A}}$, $Uz = 0$ are a generating set of $I(U)$ as a k-vector space.
- As for each cycle ω we have $Uz(\omega) = 0$, the cycle vector $z(\omega)$ belongs to $\ker_{\mathbb{Z}} U = \left\{z \in \mathbb{Z}^{\mathcal{A}} : Uz = 0\right\}$. Moreover, $P^{z^+(\omega)} = P^{\omega}$, $P^{z^-(\omega)} = P^{r\omega}$, therefore the K-ideal is contained in the toric ideal $I(U)$.

The K-ideal is toric

Theorem

The K-ideal is the toric ideal of the cocycle matrix.

Definition (Graver basis)

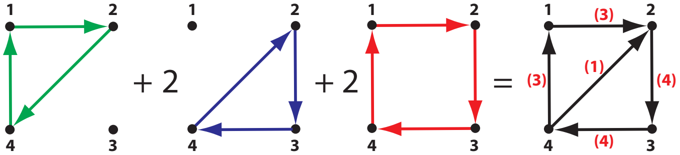
$z(\omega_1)$ is conformal to $z(\omega_2)$, $z(\omega_1) \sqsubseteq z(\omega_2)$, if the component-wise product is non-negative and $|z(\omega_1)| \leq |z(\omega_2)|$ component-wise, i.e. $z_a(\omega_1)z_a(\omega_2) \geq 0$ and $|z_a(\omega_1)| \leq |z_a(\omega_2)|$ for all $a \in \mathcal{A}$. A Graver basis of $Z(\mathcal{O})$ is the set of the minimal elements with respect to the conformity partial order $\sqsubseteq$.

Theorem

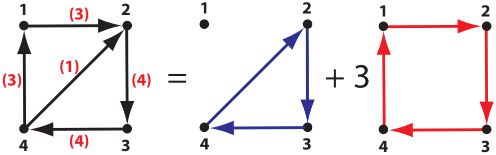
1. *For each cycle vector $z \in Z(\mathcal{O})$, $z = \sum_{\omega \in \mathcal{C}} \lambda(\omega) z(\omega)$, there exist cycles $\omega_1, \ldots, \omega_n \in \mathcal{C}$ and positive integers $\alpha(\omega_1), \ldots, \alpha(\omega_n)$, such that $z^+ \geq z^+(\omega_i)$, $z^- \geq z^-(\omega_i)$, $i = 1, \ldots, n$ and $z = \sum_{i=1}^{n} \alpha(\omega_i) z(\omega_i)$.*
2. *The set $\{z(\omega) \colon \omega \in \mathcal{C}\}$ is a Graver basis of $\mathcal{Z}(\mathcal{O})$. The binomials of the cycles form a Graver basis of the K-ideal.*

Example of proof

	$1\to2$	$2\to1$	$2\to3$	$3\to2$	$3\to4$	$4\to3$	$4\to1$	$1\to4$	$2\to4$	$4\to2$
$z(\omega_A) =$	1	−1	0	0	0	0	1	−1	1	−1
$z(\omega_B) =$	0	0	1	−1	1	−1	0	0	−1	1
$z(\omega_C) =$	1	−1	1	−1	1	−1	1	−1	0	0



$$z(\omega) = z(\omega_A) + 2z(\omega_B) + 2z(\omega_C) = (3, -3, 4, -4, 4, -4, 0, 0, -1, 1)$$
$$z^+(\omega) = z^+(\omega_B) + 3z^+(\omega_C) = (3, 0, 4, 0, 4, 0, 0, 0, 0, 1)$$



Positive K-ideal

- The strictly positive reversible transition probabilities on $\mathcal{O}$ are given by:

$$
\begin{aligned}
P_{v\to w} &= s(v,w)\prod_{S} t_S^{u_{v\to w}(S)} \\
&= s(v,w)\prod_{S:\ v\in S, w\notin S} t_S \prod_{S:\ w\in S, v\notin S} t_S^{-1},
\end{aligned}
$$

  where $s(v,w)=s(w,v)>0$, $t_S>0$.

- The first set of parameters, $s(v,w)$, is a function of the edge.
- The second set of parameters, t_S, represent the deviation from symmetry. The second set of parameters is not identifiable because the rows of the U matrix are not linearly independent.
- The parametrization can be used to derive an explicit form of the invariant probability.

Parametric detailed balance

Theorem

Consider the strictly non-zero points on the K-variety.

1. *The symmetric parameters $s(e)$, $e \in \mathcal{E}$, are uniquely determined. The parameters t_S, $S \subset V$ are confounded by* $\ker U = \{U^t t = 0\}$.
2. *An identifiable parametrization is obtained by taking a subset of parameters corresponding to linearly independent rows, denoted by t_S, $S \subset \mathcal{S}$:*

$$P_{v \to w} = s(v, w) \prod_{S \subset \mathcal{S} \colon v \in S, w \notin S} t_S \prod_{S \subset \mathcal{S} \colon w \in S, v \notin S} t_S^{-1}$$

3. *The detailed balance equations, $\kappa(v) P_{v \to w} = \kappa(w) P_{w \to v}$, are verified if, and only if,*

$$\kappa(v) \propto \prod_{S \colon v \in S} t_S^{-2}$$

Detailed balance ideal

Definition

The detailed balance ideal is the ideal

$$\text{Ideal}\left(\prod_{v\in V}\kappa(v)-1, \kappa(v)P_{v\to w}-\kappa(w)P_{v\to w},\ (v\to w)\in\mathcal{A}\right).$$

in $k[\kappa(v) : v\in V, P_{v\to w}, (v\to w)\in\mathcal{A}]$

1. The matrix $[P_{v\to w}]_{v\to w\in\mathcal{A}}$ is a point of the variety of the K-ideal if and only if there exists $\kappa=(\kappa(v): v\in V)$ such that (κ, P) belongs to the variety of the detailed balance ideal.
2. The detailed balance ideal is a toric ideal.
3. The K-ideal is the κ-elimination ideal of the detailed balance ideal.

Parameterization of reversible transitions

- There exist a (non algebraic) parametrization of the non-zero K-variety of the form
$$P_{v\to w} = s(v,w)\kappa(w)^{1/2}\kappa(v)^{-1/2}$$
- Such a P is a reversible transition probability strictly positive on the graph $\mathcal{G}$ with invariant probability proportional to κ if, and only if,
$$\kappa(v)^{1/2} \geq \sum_{w\neq v} s(u,w)\kappa(w)^{-1/2}.$$
- In the Hastings-Metropolis algorithm, we are given an unnormalized positive probability κ and a transition $Q_{v\to w} > 0$ if $(v \to w) \in \mathcal{A}$. We are required to produce a new transition $P_{v\to w} = Q_{v\to w}\alpha(v,w)$ such that P is reversible with invariant probability κ and $0 < \alpha(v,w) \leq 1$. We have
$$Q_{v\to w}\alpha(v,w) = s(v,w)\kappa(w)^{1/2}\kappa(v)^{-1/2}$$
and moreover we want
$$\alpha(v,w) = \frac{s(v,w)\kappa(w)^{1/2}}{Q_{v\to w}\kappa(v)^{1/2}} \leq 1.$$