A Markov chain for lattice polytopes

GASCom 2018

Rado Rakotonarivo

LIPN, Université Paris 13, Villetaneuse, France

19 juin 2018

Joint work with Julien David and Lionel Pournin

Outlines

Lattice Polytopes

Polytope

A polytope is a convex hull of a finite set of points in an euclidean space.

Lattice Polytopes

Polytope

A polytope is a convex hull of a finite set of points in an euclidean space.

Lattice Polytope

A lattice polytope is a polytope such that its vertices have integer coordinates.

Lattice Polytopes

Polytope

A polytope is a convex hull of a finite set of points in an euclidean space.

Lattice Polytope

A lattice polytope is a polytope such that its vertices have integer coordinates.

Full dimension Lattice Polytope

A lattice polytope is called full-dimensional if it is a d-dimensional object in $\mathbb{N}^d$.

Lattice Polytopes

Polytope

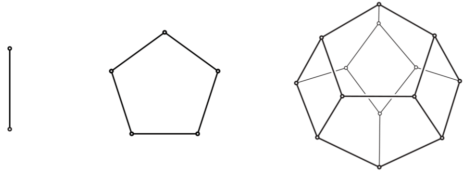
A polytope is a convex hull of a finite set of points in an euclidean space.

Lattice Polytope

A lattice polytope is a polytope such that its vertices have integer coordinates.

Full dimension Lattice Polytope

A lattice polytope is called full-dimensional if it is a d-dimensional object in $\mathbb{N}^d$.



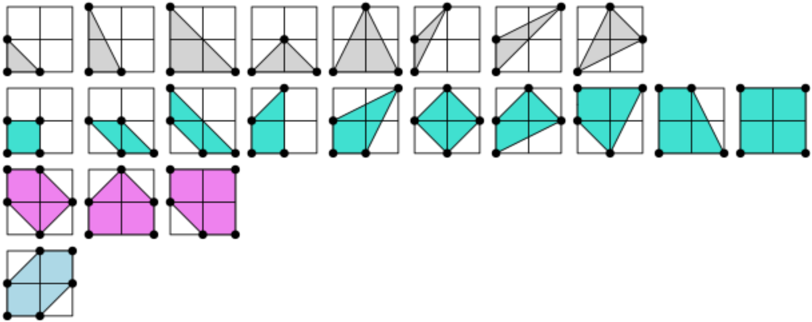
Lattice (d, k)-polytope

- A lattice (d, k)-polytope is a lattice polytope contained in a d-dimensional hypercube with side of length k.

Lattice (d, k)-polytope

- A lattice (d, k)-polytope is a lattice polytope contained in a d-dimensional hypercube with side of length k.

Example $d = 2, k = 2$



Not exhaustive list !!

Motivations

Lattice (d, k)-polytopes :

- appear in both theoritical and applied mathematical fields.
- have been studied on their properties as a function of d and k.
- have a central place in *linear optimization*.

Motivations

Lattice (d, k)-polytopes :

- appear in both theoritical and applied mathematical fields.
- have been studied on their properties as a function of d and k.
- have a central place in *linear optimization*.

However :

- practically few things are known in high dimension : enumeration ? average number of vertices ? ...
- exhaustive enumeration is prohibited due to combinatoric explosion.

Random sampling

Why ?

- to investigate the average properties of large-sized objects.
- to study the average behavior of algorithms applied to them.
- a concrete statistic analysis on abstract structures.

Random sampling

Why ?

- to investigate the average properties of large-sized objects.
- to study the average behavior of algorithms applied to them.
- a concrete statistic analysis on abstract structures.

How ?

- Ad-hoc methods.
- Combinatoric approach : recursive methods, Boltzmann samplers ...
- Probabilistic approach : Markov chains.

Random sampling

For fixed values of d and k, we are interested in the uniform distribution over all the lattice (d, k)-polytopes.

Random sampling

For fixed values of d and k, we are interested in the uniform distribution over all the lattice (d, k)-polytopes.

What has been done ?

Random samplers when $d = 2$.

- Random sampler of convex polygons in a disc [Devillers, Duchon, Thomasse '14]
- Boltzmann sampler of convex polyominoes [Bodini, Duchon, Jacquot, Mutafchiev '13]

Random sampling

For fixed values of d and k, we are interested in the uniform distribution over all the lattice (d, k)-polytopes.

What has been done?

Random samplers when $d = 2$.

- Random sampler of convex polygons in a disc [Devillers, Duchon, Thomasse '14]
- Boltzmann sampler of convex polyominoes [Bodini, Duchon, Jacquot, Mutafchiev '13]

Our contribution

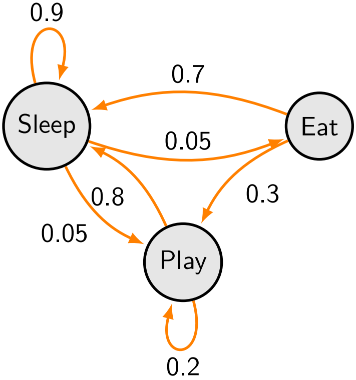
Uniform random sampler in general dimension using Markov chain.

Reminders on Markov Chains

What is Doudou doing ?

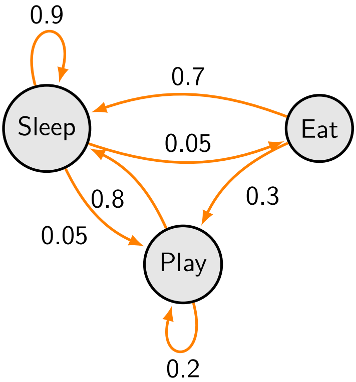
Reminders on Markov Chains

What is Doudou doing ?



Reminders on Markov Chains

What is Doudou doing ?



Example

The space of states
$\Omega = \{Sleep, Eat, Play\}$

(P)

	Sleep	Eat	Play
Sleep	0.9	0.05	0.05
Eat	0.7	0	0.3
Play	0.8	0	0.2

Stationnary distribution

- Let P be the transition matrix which describes the transition rules over Ω.
- A stationnary distribution π is a distribution witch satisfies : $\pi = \pi P$.

Stationnary distribution

- Let P be the transition matrix which describes the transition rules over Ω.
- A stationnary distribution π is a distribution witch satisfies : $\pi = \pi P$.

Example

A stationnary distribution for Doudou is :

$$\pi = [0.884 \quad 0.0442 \quad 0.0718].$$

Doudou spents 88.4% of his time sleeping !!

- Observing Doudou when it reaches its stationnary distribution means we pick at random a state following the distribution π.

From a Markov chain to a random sampler

- We are interested in the uniform distribution over Ω.

From a Markov chain to a random sampler

- We are interested in the uniform distribution over Ω.
- Some conditions on the Markov chain garanties that the uniform distribution is the unique stationnary distribution :

From a Markov chain to a random sampler

- We are interested in the uniform distribution over Ω.
- Some conditions on the Markov chain garanties that the uniform distribution is the unique stationnary distribution :

Irreducibility
- The underlying graph of the Markov chain is connected.
- Every state is reacheable from any state in a finite number of steps.

From a Markov chain to a random sampler

- We are interested in the uniform distribution over Ω.
- Some conditions on the Markov chain garanties that the uniform distribution is the unique stationnary distribution :

Irreducibility
- The underlying graph of the Markov chain is connected.
- Every state is reacheable from any state in a finite number of steps.

Aperiodicity
- All the states has a period 1, where the period of a state x is the $\gcd$(return times on x).
- Note that an for an irreducible chain, every state has the same period.

From a Markov chain to a random sampler

- We are interested in the uniform distribution over Ω.
- Some conditions on the Markov chain garanties that the uniform distribution is the unique stationnary distribution :

Irreducibility
- The underlying graph of the Markov chain is connected.
- Every state is reacheable from any state in a finite number of steps.

Aperiodicity
- All the states has a period 1, where the period of a state x is the $\gcd$(return times on x).
- Note that an for an irreducible chain, every state has the same period.

Symmetry For any two states x and $y \in \Omega$, $P(x, y) = P(y, x)$.

From a Markov chain to a random sampler

Principle

1. Build a Markov chain which space of states Ω is the set of lattice (d, k)-polytopes.

From a Markov chain to a random sampler

Principle

1. Build a Markov chain which space of states Ω is the set of lattice (d, k)-polytopes.
2. Set up a set of transition rules (local operations) which permits to move along the chain and such that the stationnary distribution is the uniform.

From a Markov chain to a random sampler

Principle

1. Build a Markov chain which space of states Ω is the set of lattice (d, k)-polytopes.
2. Set up a set of transition rules (local operations) which permits to move along the chain and such that the stationnary distribution is the uniform.
3. Run a long enough random walk on the chain until we are **close enough** to the stationnary distribution.

From a Markov chain to a random sampler

Principle

1. Build a Markov chain which space of states Ω is the set of lattice (d, k)-polytopes.
2. Set up a set of transition rules (local operations) which permits to move along the chain and such that the stationnary distribution is the uniform.
3. Run a long enough random walk on the chain until we are **close enough** to the stationnary distribution.
4. End up the walk in a (d, k)-polytope.

From a Markov chain to a random sampler

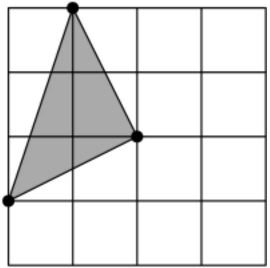
Principle

1. Build a Markov chain which space of states Ω is the set of lattice (d, k)-polytopes.
2. Set up a set of transition rules (local operations) which permits to move along the chain and such that the stationnary distribution is the uniform.
3. Run a long enough random walk on the chain until we are **close enough** to the stationnary distribution.
4. End up the walk in a (d, k)-polytope.

We want to define the simplest set of transition rules, in order to later facilitate proofs on the necessary length of the walk.

A Markov chain for (d, k)-polytopes

$\Omega :=$ set of lattice (d, k)-polytopes.
Let $P \in \Omega$ be a (d, k)-polytope with vertex set $\mathcal{V}$.

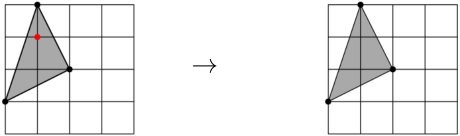


Transition rules

- sample at random a point $x \in [0, k]^d$.

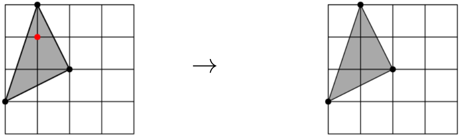
Transition rules

- sample at random a point $x \in [0, k]^d$.
- If x is contained in P but is not a vertex of it, we loop on P.

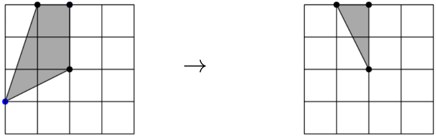


Transition rules

- sample at random a point $x \in [0,k]^d$.
- If x is contained in P but is not a vertex of it, we loop on P.

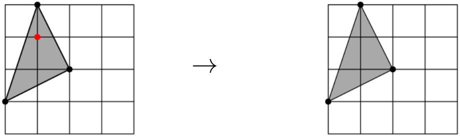


- If x is a vertex of P, then
  - If $Q = \mathrm{conv}(\mathcal{V}\backslash\{x\})$ is d-dimensional, the transition goes from P to Q.

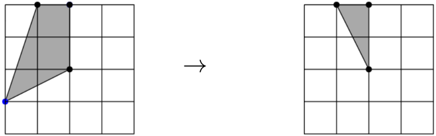


Transition rules

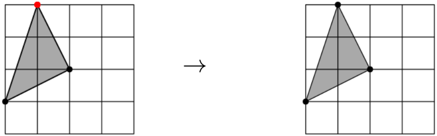
- sample at random a point $x \in [0, k]^d$.
- If x is contained in P but is not a vertex of it, we loop on P.



- If x is a vertex of P, then
  - If $Q = \mathrm{conv}(\mathcal{V} \backslash \{x\})$ is d-dimensional, the transition goes from P to Q.



  - Otherwise, we loop on P.

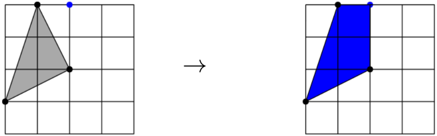


Transition rules

- If x does not belong to P, then

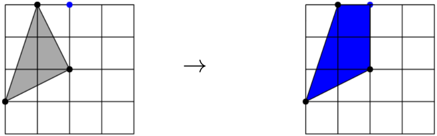
Transition rules

- If x does not belong to P, then
  - If $\mathcal{V} \cup \{x\}$ is precisely the vertex set of its convex hull, then the transition goes from P to $Q = \mathrm{conv}(\mathcal{V} \cup \{x\})$.

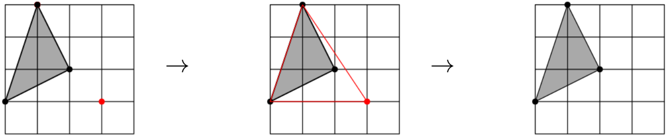


Transition rules

- If x does not belong to P, then
  - If $\mathcal{V} \cup \{x\}$ is precisely the vertex set of its convex hull, then the transition goes from P to $Q = \mathrm{conv}(\mathcal{V} \cup \{x\})$.



  - Otherwise, we loop on P.



Resulting random sampler

Theorem

For all $d \geq 2$ and for all positive k, the Markov chain is irreducible, aperiodic and symmetric.

Resulting random sampler

Theorem

For all $d \geq 2$ and for all positive k, the Markov chain is irreducible, aperiodic and symmetric.

Corollary

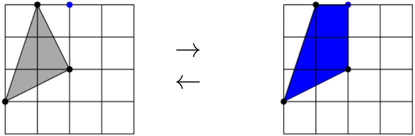
The resulting random sampler is a quasi-uniform random sampler for d-dimensional (d, k)-polytopes.

Resulting random sampler

Sketch of proof

Symmetry Transition rules ensure that for any P and $Q \in \Omega$,

$$\mathbb{P}(P, Q) = \mathbb{P}(Q, P) = \begin{cases} \frac{1}{(k+1)^d} \\ 0 \end{cases}$$

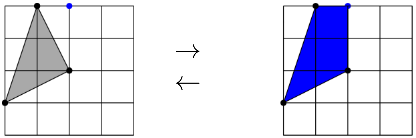


Resulting random sampler

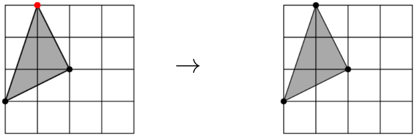
Sketch of proof

Symmetry Transition rules ensure that for any P and $Q \in \Omega$,

$$\mathbb{P}(P, Q) = \mathbb{P}(Q, P) = \begin{cases} \frac{1}{(k+1)^d} \\ 0 \end{cases}$$



Aperiodicity There always exists $P \in \Omega$ such that one can have a loop on it (take P as a simplex).



Resulting random sampler

Sketch of proof

Irreducibility 1. Any $P \in \Omega$ can be reduced as a simplex by a succession of deletion move.

Resulting random sampler

Sketch of proof

Irreducibility

1. Any $P \in \Omega$ can be reduced as a simplex by a succession of deletion move.
2. Given $P \in \Omega$, there always exists a simplex S from which after a succession of insertion move, we obtain P.

Resulting random sampler

Sketch of proof

Irreducibility

1. Any $P \in \Omega$ can be reduced as a simplex by a succession of deletion move.
2. Given $P \in \Omega$, there always exists a simplex S from which after a succession of insertion move, we obtain P.
3. Given a simplex $S \in \Omega$, we can always perform a succession of insertion and deletion to reach any simplex S'.

Mixing time

- For any positive ε, the mixing time, $t_{\mathrm{mix}}(\varepsilon)$, of a Markov chain is the amount of time the chain needs to reach a distribution whose distance to the stationnary distribution is less than ε.

Mixing time

- For any positive ε, the mixing time, $t_{\mathrm{mix}}(\varepsilon)$, of a Markov chain is the amount of time the chain needs to reach a distribution whose distance to the stationnary distribution is less than ε.
- The diameter δ, of the Markov chain's diameter is a lower bound on the mixing time.

Mixing time

- For any positive ε, the mixing time, $t_{\mathrm{mix}}(\varepsilon)$, of a Markov chain is the amount of time the chain needs to reach a distribution whose distance to the stationnary distribution is less than ε.
- The diameter δ, of the Markov chain's diameter is a lower bound on the mixing time.

An intuitive lower bound

$$\delta \geq \underbrace{n_{\mathrm{max}}}_{\text{largest number of vertices}} - \underbrace{(d+1)}_{\text{number of vertices of a simplex}} .$$

Lower bounds on mixing time

dimension	$n_{\max}$	lower bound on t_{mix}
$d = 2$	$12\left(\frac{k}{2\pi}\right)^{2/3} + O(k^{1/3}\log k)$ (1)	$t_{\text{mix}} \geq ck^{2/3}$
$d \geq 2$	$c_1(d)r^{d\frac{d-1}{d+1}}$ (2)	$t_{\text{mix}} \geq ck^{d\frac{d-1}{d+1}}$

(1) Largest number of vertices of a polygon in $[0, k]^2$ [Acketa, Zunic '95, Deza, Manoussakis, Onn '18].

(2) Largest number of faces of each dimension of a lattice polytope contained in a d-dimensional disc of radius r and centered in 0 [Barany, Larman '98].

Experimental results in dimension $d = 2$

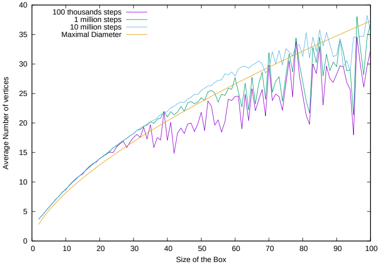


FIGURE – Number of vertices

$$\mathbb{E}[n] \geq 6 \left(\frac{k}{2\pi}\right)^{2/3} \sim \frac{n_{\max}}{2}$$

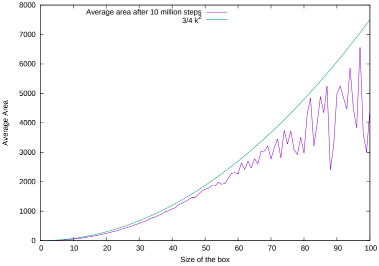


FIGURE – Area

$$\mathbb{E}[a] \leq \frac{3}{4} k^2$$

Conclusion

- Uniform distribution over the (d, k)-polytopes.
- Built a random sampler resulting from a Markov chain.
- Obtain better bounds on mixing time using spectral gap analysis.
- Find out different transition rules that makes it easier to sample.

Thank you for your attention !