Infinite–Dimensional Diffusion Processes Approximated by Finite Markov Chains on Partitions

Leonid Petrov

IPPI — Institute for Information Transmission Problems (Moscow, Russia)

May 20, 2011

1. de Finetti's Theorem
2. up/down Markov chains and limiting diffusions
3. Kingman's exchangeable random partitions
4. up/down Markov chains on partitions and limiting infinite-dimensional diffusions

de Finetti's theorem

Exchangeable binary sequences

- $\xi_1, \xi_2, \xi_3, \ldots$ — random variables $\in \{0, 1\}$

Exchangeable binary sequences

- $\xi_1, \xi_2, \xi_3, \ldots$ — random variables $\in \{0, 1\}$
- distribution of the sequence $\xi = (\xi_1, \xi_2, \xi_3, \ldots)$ is invariant under pemutations

Exchangeable binary sequences

- $\xi_1, \xi_2, \xi_3, \ldots$ — random variables $\in \{0, 1\}$
- distribution of the sequence $\xi = (\xi_1, \xi_2, \xi_3, \ldots)$ is invariant under pemutations

de Finetti's Theorem

$$\text{Exchangeable binary sequences } \xi$$
$$\updownarrow$$
$$\text{probability measures } P \text{ on } [0, 1]$$

From ξ to P

$$\frac{\#\{\xi_i = 1\colon 1 \le i \le n\}}{n} \xrightarrow{Law} P, \qquad n \to \infty$$

distributions on $\{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1\}$ approximate the measure P on $[0, 1]$.

From ξ to P

$$\frac{\#\{\xi_i = 1 \colon 1 \leq i \leq n\}}{n} \xrightarrow{Law} P, \qquad n \to \infty$$

distributions on $\{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1\}$ approximate the measure P on $[0, 1]$.

From P to ξ

1. Sample real number $p \in [0, 1]$ according to P

From ξ to P

$$\frac{\#\{\xi_i = 1 \colon 1 \le i \le n\}}{n} \xrightarrow{Law} P, \qquad n \to \infty$$

distributions on $\{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\}$ approximate the measure P on $[0,1]$.

From P to ξ

1. Sample real number $p \in [0,1]$ according to P
2. Sample Bernoulli sequence $\xi = (\xi_1, \xi_2, \dots)$ with probability p of success.

$$P_n(k) := \mathrm{Prob}(\#\{\xi_i = 1 \colon 1 \le i \le n\} = k) = \binom{n}{k} \int_0^1 p^k (1-p)^{n-k} P(dp).$$

de Finetti's theorem

- Bernoulli sequences are the *extremal* exchangeable binary sequences (P on $[0, 1]$ is δ_p)

- Bernoulli sequences are the *extremal* exchangeable binary sequences (P on $[0, 1]$ is δ_p)
- Any exchangeable binary sequence is a *mixture* of Bernoulli sequences.

finite level distributions

- P_n on $\{0, 1, 2, \ldots, n-1, n\}$

$$P_n(k) = \text{Prob}(\#\{\xi_i = 1 \colon 1 \leq i \leq n\} = k).$$

(floors of the Pascal triangle)

finite level distributions

- P_n on $\{0, 1, 2, \ldots, n-1, n\}$
$$P_n(k) = \text{Prob}(\#\{\xi_i = 1 \colon 1 \leq i \leq n\} = k).$$
(floors of the Pascal triangle)
- The distributions $\{P_n\}_{n=1}^{\infty}$ *encode* the exchangeable binary sequence ξ. The scaling limit of P_n's is our measure P on $[0, 1]$.

finite level distributions

- P_n on $\{0, 1, 2, \dots, n-1, n\}$

$$P_n(k) = \text{Prob}(\#\{\xi_i = 1 \colon 1 \le i \le n\} = k).$$

  (floors of the Pascal triangle)
- The distributions $\{P_n\}_{n=1}^{\infty}$ *encode* the exchangeable binary sequence ξ. The scaling limit of P_n's is our measure P on $[0, 1]$.
- The P_n's are *compatible* with each other:

$$P_n(k) = \tfrac{n+1-k}{n+1} P_{n+1}(k) + \tfrac{k+1}{n+1} P_{n+1}(k+1) \qquad \forall k$$

- P_n on $\{0, 1, 2, \ldots, n-1, n\}$
$$P_n(k) = \text{Prob}(\#\{\xi_i = 1 \colon 1 \le i \le n\} = k).$$
(floors of the Pascal triangle)
- The distributions $\{P_n\}_{n=1}^{\infty}$ *encode* the exchangeable binary sequence ξ. The scaling limit of P_n's is our measure P on $[0, 1]$.
- The P_n's are *compatible* with each other:
$$P_n(k) = \tfrac{n+1-k}{n+1} P_{n+1}(k) + \tfrac{k+1}{n+1} P_{n+1}(k+1) \qquad \forall k$$
$$\Updownarrow$$
Can define *down kernel* $p^{\downarrow}_{n+1,n}$ from $\{0, \ldots, n+1\}$ to $\{0, \ldots, n\}$ such that
$$P_{n+1} \circ p^{\downarrow}_{n+1,n} = P_n \qquad \forall n$$

down transition kernel

$$p^{\downarrow}_{n+1,n}(k,k) = \frac{n+k-1}{n+1}, \qquad p^{\downarrow}_{n+1,n}(k,k-1) = \frac{k}{n+1}.$$

down transition kernel

$$p^{\downarrow}_{n+1,n}(k,k) = \frac{n+k-1}{n+1}, \qquad p^{\downarrow}_{n+1,n}(k,k-1) = \frac{k}{n+1}.$$

1. Markov transition kernel: $\sum_m p^{\downarrow}_{n+1,n}(k,m) = 1$ for all k.

down transition kernel

$$p^{\downarrow}_{n+1,n}(k,k) = \frac{n+k-1}{n+1}, \qquad p^{\downarrow}_{n+1,n}(k,k-1) = \frac{k}{n+1}.$$

1. Markov transition kernel: $\sum_m p^{\downarrow}_{n+1,n}(k,m) = 1$ for all k.
2. Can define through the Pascal triangle:

$$p^{\downarrow}_{n+1,n}(k,m) = \frac{\#\text{ of paths to }(n,m)}{\#\text{ of paths to }(n+1,k)}.$$

down transition kernel

$$p^{\downarrow}_{n+1,n}(k,k) = \frac{n+k-1}{n+1}, \qquad p^{\downarrow}_{n+1,n}(k,k-1) = \frac{k}{n+1}.$$

1. Markov transition kernel: $\sum_m p^{\downarrow}_{n+1,n}(k,m) = 1$ for all k.
2. Can define through the Pascal triangle:

$$p^{\downarrow}_{n+1,n}(k,m) = \frac{\#\text{ of paths to }(n,m)}{\#\text{ of paths to }(n+1,k)}.$$

summarizing

- Binary sequence ξ is exchangeable iff $\{P_n\}$'s are compatible with the down kernel: $P_{n+1} \circ p^{\downarrow}_{n+1,n} = P_n$.

down transition kernel

$$p^{\downarrow}_{n+1,n}(k,k) = \frac{n+k-1}{n+1}, \qquad p^{\downarrow}_{n+1,n}(k,k-1) = \frac{k}{n+1}.$$

1. Markov transition kernel: $\sum_m p^{\downarrow}_{n+1,n}(k,m) = 1$ for all k.
2. Can define through the Pascal triangle:

$$p^{\downarrow}_{n+1,n}(k,m) = \frac{\#\text{ of paths to }(n,m)}{\#\text{ of paths to }(n+1,k)}.$$

summarizing

- Binary sequence ξ is exchangeable iff $\{P_n\}$'s are compatible with the down kernel: $P_{n+1} \circ p^{\downarrow}_{n+1,n} = P_n$.
- de Finetti's Theorem = classification of compatible (*coherent*) sequences of measures $\{P_n\}$ on levels of the Pascal triangle.

- Fix a measure P on $[0, 1]$ $\Leftrightarrow$ a system $\{P_n\}$.

up transition kernel

- Fix a measure P on $[0,1] \Leftrightarrow$ a system $\{P_n\}$.
- Define

$$p^{\uparrow}_{n,n+1}(m,k) := p^{\downarrow}_{n+1,n}(k,m)\frac{P_{n+1}(k)}{P_n(m)}.$$

- Fix a measure P on $[0,1] \Leftrightarrow$ a system $\{P_n\}$.
- Define

$$p^{\uparrow}_{n,n+1}(m,k) := p^{\downarrow}_{n+1,n}(k,m)\frac{P_{n+1}(k)}{P_n(m)}.$$

Thus

$$P_n \circ p^{\uparrow}_{n,n+1} = P_{n+1} \qquad \forall n.$$

(random growth on the Pascal triangle; depends on P).

- Fix a measure P on $[0,1] \Leftrightarrow$ a system $\{P_n\}$.
- Define
$$p^{\uparrow}_{n,n+1}(m,k) := p^{\downarrow}_{n+1,n}(k,m)\frac{P_{n+1}(k)}{P_n(m)}.$$
Thus
$$P_n \circ p^{\uparrow}_{n,n+1} = P_{n+1} \qquad \forall n.$$
(random growth on the Pascal triangle; depends on P).

This is a restatement of the two-step sampling procedure for the exchangeable binary sequence $\xi = (\xi_1, \xi_2, \dots)$ corresponding to P.

1. de Finetti's Theorem
2. up/down Markov chains and limiting diffusions
3. Kingman's exchangeable random partitions
4. up/down Markov chains on partitions and limiting infinite-dimensional diffusions

up/down Markov chains

- Fix a measure P on $[0, 1]$.

up/down Markov chains

- Fix a measure P on $[0, 1]$.
- Thus get the up kernel $p^{\uparrow}_{n,n+1}$.

up/down Markov chains

- Fix a measure P on $[0, 1]$.
- Thus get the up kernel $p^{\uparrow}_{n,n+1}$.
- Also there is a canonical down kernel $p^{\downarrow}_{n+1,n}$.

up/down Markov chains

- Fix a measure P on $[0, 1]$.
- Thus get the up kernel $p^{\uparrow}_{n,n+1}$.
- Also there is a canonical down kernel $p^{\downarrow}_{n+1,n}$.
- Use these kernels to obtain the *up/down Markov chain* on $\{0, 1, \ldots, n\}$:

$$T_n = p^{\uparrow}_{n,n+1} \circ p^{\downarrow}_{n+1,n},$$

i.e.,

$$T_n(k, \tilde{k}) = \sum_{m=0}^{n+1} p^{\uparrow}_{n,n+1}(k, m) \circ p^{\downarrow}_{n+1,n}(m, \tilde{k}).$$

up/down Markov chains

- Fix a measure P on $[0,1]$.
- Thus get the up kernel $p^{\uparrow}_{n,n+1}$.
- Also there is a canonical down kernel $p^{\downarrow}_{n+1,n}$.
- Use these kernels to obtain the *up/down Markov chain* on $\{0,1,\ldots,n\}$:

$$T_n = p^{\uparrow}_{n,n+1} \circ p^{\downarrow}_{n+1,n},$$

  i.e.,

$$T_n(k,\tilde{k}) = \sum_{m=0}^{n+1} p^{\uparrow}_{n,n+1}(k,m) \circ p^{\downarrow}_{n+1,n}(m,\tilde{k}).$$

- P_n is reversible w.r.t. T_n for all n.

up/down Markov chains

- Fix a measure P on $[0, 1]$.
- Thus get the up kernel $p^{\uparrow}_{n,n+1}$.
- Also there is a canonical down kernel $p^{\downarrow}_{n+1,n}$.
- Use these kernels to obtain the *up/down Markov chain* on $\{0, 1, \ldots, n\}$:

$$T_n = p^{\uparrow}_{n,n+1} \circ p^{\downarrow}_{n+1,n},$$

i.e.,

$$T_n(k, \tilde{k}) = \sum_{m=0}^{n+1} p^{\uparrow}_{n,n+1}(k, m) \circ p^{\downarrow}_{n+1,n}(m, \tilde{k}).$$

- P_n is reversible w.r.t. T_n for all n.

Is there a (scaling) limit of the chains T_n as $n \to \infty$?

distinguished measures P on $[0, 1]$

Is there a (scaling) limit of the chains T_n as $n \to \infty$?

distinguished measures P on $[0, 1]$

Is there a (scaling) limit of the chains T_n as $n \to \infty$?

$\to$ Nothing general can be said for arbitrary P on $[0, 1]$.

distinguished measures P on $[0,1]$

Is there a (scaling) limit of the chains T_n as $n \to \infty$?
$\to$ Nothing general can be said for arbitrary P on $[0,1]$.

Beta distributions

Beta distributions $P_{a,b}$ (where $a, b > 0$):

$$P_{a,b}(dx) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}dx.$$

distinguished measures P on $[0, 1]$

Is there a (scaling) limit of the chains T_n as $n \to \infty$?
$\to$ Nothing general can be said for arbitrary P on $[0, 1]$.

Beta distributions

Beta distributions $P_{a,b}$ (where $a, b > 0$):

$$P_{a,b}(dx) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} dx.$$

- One gets diffusion limits for the chains T_n

distinguished measures P on $[0,1]$

Is there a (scaling) limit of the chains T_n as $n \to \infty$?
$\to$ Nothing general can be said for arbitrary P on $[0,1]$.

Beta distributions

Beta distributions $P_{a,b}$ (where $a, b > 0$):

$$P_{a,b}(dx) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} dx.$$

- One gets diffusion limits for the chains T_n
- $P_{a,b}$'s are nice from algebraic point

distinguished measures P on $[0,1]$

Is there a (scaling) limit of the chains T_n as $n \to \infty$?
$\to$ Nothing general can be said for arbitrary P on $[0,1]$.

Beta distributions

Beta distributions $P_{a,b}$ (where $a, b > 0$):

$$P_{a,b}(dx) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}dx.$$

- One gets diffusion limits for the chains T_n
- $P_{a,b}$'s are nice from algebraic point
- *Remark:* $P_{a,b}$'s from a Bayesian point:

  Beta prior on Bernoulli parameter $p \in [0,1]$
  $\Rightarrow$ Beta posterior

Beta distributions and the uniform distribution

Take $P_{a,b}$ on $[0,1]$, then

$$P_n(k) = \int_0^1 \frac{\Gamma(n+1)\Gamma(a+b)}{\Gamma(k+1)\Gamma(n-k+1)\Gamma(a)\Gamma(b)} p^{k+a-1}(1-p)^{n-k+b-1} dp$$

$$= \binom{n}{k} \frac{(a)_k (b)_{n-k}}{(a+b)_n},$$

where

$$(x)_m := x(x+1)(x+2)\dots(x+m-1).$$

Beta distributions and the uniform distribution

Take $P_{a,b}$ on $[0,1]$, then

$$P_n(k) = \int_0^1 \frac{\Gamma(n+1)\Gamma(a+b)}{\Gamma(k+1)\Gamma(n-k+1)\Gamma(a)\Gamma(b)} p^{k+a-1}(1-p)^{n-k+b-1} dp$$

$$= \binom{n}{k} \frac{(a)_k (b)_{n-k}}{(a+b)_n},$$

where

$$(x)_m := x(x+1)(x+2)\dots(x+m-1).$$

Uniform distribution

For the uniform distribution ($a = b = 1$) on $[0,1]$, P_n is also uniform on $\{0, 1, \dots, n\}$.

up/down Markov chains for uniform distribution

- $p^{\uparrow}_{n,n+1}(k, k+1) = \dfrac{k+1}{n+2}, \qquad p^{\uparrow}_{n,n+1}(k, k) = \dfrac{n-k+1}{n+2}.$

up/down Markov chains for uniform distribution

- $p^{\uparrow}_{n,n+1}(k, k+1) = \dfrac{k+1}{n+2}, \qquad p^{\uparrow}_{n,n+1}(k, k) = \dfrac{n-k+1}{n+2}.$
- $T_n(k \to k+1) = \frac{(k+1)(n-k)}{(n+1)(n+2)}$

  $T_n(k \to k-1) = \frac{k(n-k+1)}{(n+1)(n+2)}$

  $T_n(k \to k) = \frac{(k+1)^2+(n-k+1)^2}{(n+1)(n+2)}$

up/down Markov chains for uniform distribution

- $p^{\uparrow}_{n,n+1}(k, k+1) = \dfrac{k+1}{n+2}, \qquad p^{\uparrow}_{n,n+1}(k, k) = \dfrac{n-k+1}{n+2}.$
- $T_n(k \to k+1) = \frac{(k+1)(n-k)}{(n+1)(n+2)}$

  $T_n(k \to k-1) = \frac{k(n-k+1)}{(n+1)(n+2)}$

  $T_n(k \to k) = \frac{(k+1)^2+(n-k+1)^2}{(n+1)(n+2)}$
- $\dfrac{k}{n} = x \in [0,1],\ \Delta x = \dfrac{1}{n}$:

  $T_n(k \to k+1) = \text{Prob}(x \to x+\Delta x) \sim x(1-x) + \frac{1}{n}(1-x)$

  $T_n(k \to k-1) = \text{Prob}(x \to x-\Delta x) \sim x(1-x) + \frac{1}{n}x$

up/down Markov chains for uniform distribution

Scaling limit

- Scale space by $\frac{1}{n}$, scale time by $\frac{1}{n^2}$ (i.e., one step of the nth Markov chain = small time interval of order n^{-2}).

up/down Markov chains for uniform distribution

Scaling limit

- Scale space by $\frac{1}{n}$, scale time by $\frac{1}{n^2}$ (i.e., one step of the nth Markov chain = small time interval of order n^{-2}).
- Then as $n \to +\infty$,

$$n^2(T_n - \mathbf{1}) \to x(1-x)\frac{d^2}{dx^2} + (1-2x)\frac{d}{dx}$$

up/down Markov chains for uniform distribution

Scaling limit

- Scale space by $\frac{1}{n}$, scale time by $\frac{1}{n^2}$ (i.e., one step of the nth Markov chain = small time interval of order n^{-2}).
- Then as $n \to +\infty$,

$$n^2(T_n - \mathbf{1}) \to x(1-x)\frac{d^2}{dx^2} + (1-2x)\frac{d}{dx}$$

- First, check the convergence on polynomials (*algebraically!*). The processes have polynomial core.

up/down Markov chains for uniform distribution

Scaling limit

- Scale space by $\frac{1}{n}$, scale time by $\frac{1}{n^2}$ (i.e., one step of the nth Markov chain = small time interval of order n^{-2}).
- Then as $n \to +\infty$,

$$n^2(T_n - \mathbf{1}) \to x(1-x)\frac{d^2}{dx^2} + (1-2x)\frac{d}{dx}$$

- First, check the convergence on polynomials (*algebraically!*). The processes have polynomial core.
- Then apply Trotter-Kurtz type theorems to conclude the convergence of finite Markov chains to diffusion processes.

1. de Finetti's Theorem
2. up/down Markov chains and limiting diffusions

intermission: a variety of models

3. Kingman's exchangeable random partitions
4. up/down Markov chains on partitions and limiting infinite-dimensional diffusions

intermission: a variety of models

Counting paths $\Rightarrow p^{\downarrow} \Rightarrow$ classification of coherent measures

intermission: a variety of models

Counting paths $\Rightarrow p^{\downarrow} \Rightarrow$ classification of coherent measures

Examples

- Pascal triangle

intermission: a variety of models

Counting paths $\Rightarrow p^{\downarrow} \Rightarrow$ classification of coherent measures

Examples

- Pascal triangle
- Kingman's exchangeable random partitions

intermission: a variety of models

Counting paths $\Rightarrow p^{\downarrow} \Rightarrow$ classification of coherent measures

Examples

- Pascal triangle
- Kingman's exchangeable random partitions

- q-Pascal triangle, Stirling triangles, etc.

intermission: a variety of models

Counting paths $\Rightarrow p^{\downarrow} \Rightarrow$ classification of coherent measures

Examples

- Pascal triangle
- Kingman's exchangeable random partitions

- q-Pascal triangle, Stirling triangles, etc.
- Young graph — the lattice of all partitions

intermission: a variety of models

Counting paths $\Rightarrow p^{\downarrow} \Rightarrow$ classification of coherent measures

Examples

- Pascal triangle
- Kingman's exchangeable random partitions

- q-Pascal triangle, Stirling triangles, etc.
- Young graph — the lattice of all partitions
- Young graph with multiplicities (one choice of multiplicities is equivalent to Kingman's theory)

intermission: a variety of models

Counting paths $\Rightarrow p^{\downarrow} \Rightarrow$ classification of coherent measures

Examples

- Pascal triangle
- Kingman's exchangeable random partitions

- q-Pascal triangle, Stirling triangles, etc.
- Young graph — the lattice of all partitions
- Young graph with multiplicities (one choice of multiplicities is equivalent to Kingman's theory)
- Schur graph — the lattice of all strict partitions (*my yesterday's talk:* a related Markov dynamics is Pfaffian)

intermission: a variety of models

Counting paths $\Rightarrow p^{\downarrow} \Rightarrow$ classification of coherent measures

Examples

- Pascal triangle
- Kingman's exchangeable random partitions

- q-Pascal triangle, Stirling triangles, etc.
- Young graph — the lattice of all partitions
- Young graph with multiplicities (one choice of multiplicities is equivalent to Kingman's theory)
- Schur graph — the lattice of all strict partitions (*my yesterday's talk:* a related Markov dynamics is Pfaffian)
- Gelfand-Tsetlin schemes

...

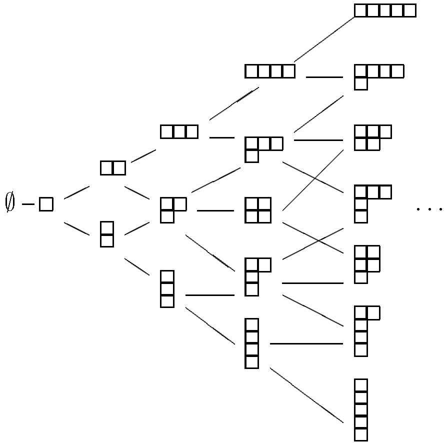

Figure: Young graph

intermission: a variety of models

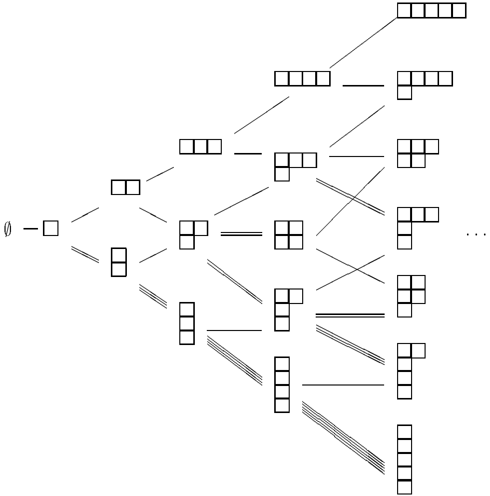

Figure: Kingman graph (= Young graph with multiplicities)

1. de Finetti's Theorem
2. up/down Markov chains and limiting diffusions
3. Kingman's exchangeable random partitions
4. up/down Markov chains on partitions and limiting infinite-dimensional diffusions

exchangeable random partitions

Definition

π is a random partition of $\mathbb{N} = \{1, 2, 3, \ldots\}$, i.e., $\mathbb{N} = \bigsqcup_{i \in I} A_i$.

Definition

π is a random partition of $\mathbb{N} = \{1, 2, 3, \dots\}$, i.e., $\mathbb{N} = \bigsqcup_{i \in I} A_i$.

π is called *exchangeable* if the distribution of π is invariant under permutations of $\mathbb{N}$.

exchangeable random partitions

Definition

π is a random partition of $\mathbb{N} = \{1, 2, 3, \dots\}$, i.e., $\mathbb{N} = \bigsqcup_{i \in I} A_i$.

π is called *exchangeable* if the distribution of π is invariant under permutations of $\mathbb{N}$.

Kingman Simplex

$$\overline{\nabla}_\infty := \left\{ (x_1 \geq x_2 \geq \dots \geq 0) : \sum x_i \leq 1 \right\}$$

exchangeable random partitions

Definition

π is a random partition of $\mathbb{N} = \{1, 2, 3, \dots\}$, i.e., $\mathbb{N} = \bigsqcup_{i \in I} A_i$.

π is called *exchangeable* if the distribution of π is invariant under permutations of $\mathbb{N}$.

Kingman Simplex

$$\overline{\nabla}_\infty := \left\{ (x_1 \geq x_2 \geq \cdots \geq 0) \colon \ \sum x_i \leq 1 \right\}$$

- Compact in coordinatewise topology.

exchangeable random partitions

Definition

π is a random partition of $\mathbb{N} = \{1, 2, 3, \dots\}$, i.e., $\mathbb{N} = \bigsqcup_{i \in I} A_i$.

π is called *exchangeable* if the distribution of π is invariant under permutations of $\mathbb{N}$.

Kingman Simplex

$$\overline{\nabla}_\infty := \left\{ (x_1 \geq x_2 \geq \cdots \geq 0) \colon \sum x_i \leq 1 \right\}$$

- Compact in coordinatewise topology.
- $x \in \overline{\nabla}_\infty \longleftrightarrow$ discrete distribution

exchangeable random partitions

Definition

π is a random partition of $\mathbb{N} = \{1, 2, 3, \dots\}$, i.e., $\mathbb{N} = \bigsqcup_{i \in I} A_i$.

π is called *exchangeable* if the distribution of π is invariant under permutations of $\mathbb{N}$.

Kingman Simplex

$$\overline{\nabla}_\infty := \left\{ (x_1 \geq x_2 \geq \cdots \geq 0) \colon \sum x_i \leq 1 \right\}$$

- Compact in coordinatewise topology.
- $x \in \overline{\nabla}_\infty \longleftrightarrow$ discrete distribution
- Probability measure on $\overline{\nabla}_\infty \longleftrightarrow$ random discrete distribution

Kingman's representation

π — exchangeable random partition of $\mathbb{N}$

$$\updownarrow$$

probability measure M on $\overline{\nabla}_\infty$ (the *boundary measure*)

exchangeable random partitions

Kingman's representation

π — exchangeable random partition of $\mathbb{N}$

$$\updownarrow$$

probability measure M on $\overline{\nabla}_\infty$ (the *boundary measure*)

Finite level random *integer partitions* — encoding of π

- Restrict π to $\{1, 2, \dots, n\} \subset \mathbb{N}$, thus get a random partition π_n of the finite set $\{1, 2, \dots, n\}$.

exchangeable random partitions

Kingman's representation

π — exchangeable random partition of $\mathbb{N}$

$\updownarrow$

probability measure M on $\overline{\nabla}_\infty$ (the *boundary measure*)

Finite level random *integer partitions* — encoding of π

- Restrict π to $\{1, 2, \dots, n\} \subset \mathbb{N}$, thus get a random partition π_n of the finite set $\{1, 2, \dots, n\}$.
- Exchangeable $\Rightarrow$ the distribution of π_n is *encoded* by the distribution of decreasing sizes of blocks

$$\lambda = (\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{\ell(\lambda)} > 0), \quad \lambda_i \in \mathbb{Z}, \quad \sum \lambda_i = n.$$

exchangeable random partitions

Kingman's representation

π — exchangeable random partition of $\mathbb{N}$

$$\updownarrow$$

probability measure M on $\overline{\nabla}_\infty$ (the *boundary measure*)

Finite level random *integer partitions* — encoding of π

- Restrict π to $\{1, 2, \ldots, n\} \subset \mathbb{N}$, thus get a random partition π_n of the finite set $\{1, 2, \ldots, n\}$.
- Exchangeable $\Rightarrow$ the distribution of π_n is *encoded* by the distribution of decreasing sizes of blocks

$$\lambda = (\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{\ell(\lambda)} > 0), \quad \lambda_i \in \mathbb{Z}, \quad \sum \lambda_i = n.$$

- $\Rightarrow$ sequence of measures M_n on the sets $\mathbb{P}_n := \{\lambda = (\lambda_1 \geq \cdots \geq \lambda_{\ell(\lambda)})$ — integer partition such that $\sum \lambda_i = n\}$

from $\{M_n\}$ to M

- Random $x_1 \geq x_2 \geq \cdots \geq 0$ are limiting values of $\lambda_1, \lambda_2, \ldots$:

$$\frac{\lambda_j}{n} \xrightarrow{Law} x_j, \qquad j = 1, 2, \ldots.$$

exchangeable random partitions

from $\{M_n\}$ to M

- Random $x_1 \geq x_2 \geq \cdots \geq 0$ are limiting values of $\lambda_1, \lambda_2, \ldots$:

$$\frac{\lambda_j}{n} \xrightarrow{Law} x_j, \qquad j = 1, 2, \ldots.$$

- The sets of partitions $\mathbb{P}_n$ approximate $\overline{\nabla}_\infty$:

$$\mathbb{P}_n \ni \lambda = (\lambda_1, \ldots, \lambda_\ell) \hookrightarrow \left(\tfrac{\lambda_1}{n}, \ldots \tfrac{\lambda_\ell}{n}, 0, 0, \ldots\right) \in \overline{\nabla}_\infty.$$

Images of M_n's weakly converge to M on $\overline{\nabla}_\infty$.

exchangeable random partitions

from $\{M_n\}$ to M

- Random $x_1 \geq x_2 \geq \cdots \geq 0$ are limiting values of $\lambda_1, \lambda_2, \ldots$:

$$\frac{\lambda_j}{n} \xrightarrow{Law} x_j, \qquad j = 1, 2, \ldots.$$

- The sets of partitions $\mathbb{P}_n$ approximate $\overline{\nabla}_\infty$:

$$\mathbb{P}_n \ni \lambda = (\lambda_1, \ldots, \lambda_\ell) \hookrightarrow \left(\tfrac{\lambda_1}{n}, \ldots \tfrac{\lambda_\ell}{n}, 0, 0, \ldots\right) \in \overline{\nabla}_\infty.$$

  Images of M_n's weakly converge to M on $\overline{\nabla}_\infty$.

From the point of random partitions π of $\mathbb{N}$ the $x_1 \geq x_2 \geq \cdots \geq 0$ are the limiting frequencies of blocks, in decreasing order.

from M to $\{M_n\}$ — two-step random sampling

1. Sample $x = (x_1, x_2, \dots) \in \overline{\nabla}_\infty$ according to M on $\overline{\nabla}_\infty$. For simplicity, let $\sum x_i = 1$

exchangeable random partitions

from M to $\{M_n\}$ — two-step random sampling

1. Sample $x = (x_1, x_2, \dots) \in \overline{\nabla}_\infty$ according to M on $\overline{\nabla}_\infty$. For simplicity, let $\sum x_i = 1$
2. Consider the measure with weights $x_1, x_2, \dots$ on $\mathbb{N}$. Sample n independent numbers $A_1, \dots, A_n \in \mathbb{N}$ according to this measure

exchangeable random partitions

from M to $\{M_n\}$ — two-step random sampling

1. Sample $x = (x_1, x_2, \dots) \in \overline{\nabla}_\infty$ according to M on $\overline{\nabla}_\infty$. For simplicity, let $\sum x_i = 1$
2. Consider the measure with weights $x_1, x_2, \dots$ on $\mathbb{N}$. Sample n independent numbers $A_1, \dots, A_n \in \mathbb{N}$ according to this measure
3. $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_\ell$ — multiplicities of $A_1, A_2, \dots, A_n$ in decreasing order:

   e.g., $(A_1, \dots, A_n) = (4, 3, 5, 1, 1, 3, 1) \to \lambda = (3, 2, 1, 1)$

   Law of $\lambda = (\lambda_1, \dots, \lambda_\ell) \in \mathbb{P}_n$ is M_n.

monomial symmetric functions

$\mu = (\mu_1, \ldots, \mu_k)$ — integer partition,

$$m_\mu(y_1, y_2, \ldots) := \sum y_{i_1}^{\mu_1} y_{i_2}^{\mu_2} \ldots y_{i_k}^{\mu_k}$$

(sum over all distinct summands).

monomial symmetric functions

$\mu = (\mu_1, \ldots, \mu_k)$ — integer partition,

$$m_\mu(y_1, y_2, \ldots) := \sum y_{i_1}^{\mu_1} y_{i_2}^{\mu_2} \ldots y_{i_k}^{\mu_k}$$

(sum over all distinct summands).
Examples:

monomial symmetric functions

$\mu = (\mu_1, \ldots, \mu_k)$ — integer partition,

$$m_\mu(y_1, y_2, \ldots) := \sum y_{i_1}^{\mu_1} y_{i_2}^{\mu_2} \ldots y_{i_k}^{\mu_k}$$

(sum over all distinct summands).
Examples:

- $m_{(k)}(y_1, y_2, \ldots) := p_k(y_1, y_2, \ldots) = \sum_{i=1}^{\infty} y_i^k$ (Newton power sum)

monomial symmetric functions

$\mu = (\mu_1, \ldots, \mu_k)$ — integer partition,

$$m_\mu(y_1, y_2, \ldots) := \sum y_{i_1}^{\mu_1} y_{i_2}^{\mu_2} \ldots y_{i_k}^{\mu_k}$$

(sum over all distinct summands).
Examples:

- $m_{(k)}(y_1, y_2, \ldots) := p_k(y_1, y_2, \ldots) = \sum_{i=1}^{\infty} y_i^k$ (Newton power sum)
- $m_{(\underbrace{1, \ldots, 1}_{k})}(y_1, y_2, \ldots) = \sum_{i_1 < \cdots < i_k} y_{i_1} \ldots y_{i_k}$ (elementary symmetric function)

exchangeable random partitions

monomial symmetric functions

$\mu = (\mu_1, \ldots, \mu_k)$ — integer partition,

$$m_\mu(y_1, y_2, \ldots) := \sum y_{i_1}^{\mu_1} y_{i_2}^{\mu_2} \ldots y_{i_k}^{\mu_k}$$

(sum over all distinct summands).
Examples:

- $m_{(k)}(y_1, y_2, \ldots) := p_k(y_1, y_2, \ldots) = \sum_{i=1}^{\infty} y_i^k$ (Newton power sum)
- $m_{(\underbrace{1, \ldots, 1}_{k})}(y_1, y_2, \ldots) = \sum_{i_1 < \cdots < i_k} y_{i_1} \ldots y_{i_k}$ (elementary symmetric function)
- For $k > 1$: $m_{(k,1)}(y_1, y_2, \ldots) = \sum_{i,j=1}^{\infty} y_i^k y_j$.

from M to $\{M_n\}$ — two-step random sampling

For simplicity, let M be concentrated on $\{\sum x_i = 1\}$.

$$M_n(\lambda) = \binom{n}{\lambda_1, \lambda_2, \ldots, \lambda_\ell} \int_{\overline{\nabla}_\infty} m_\lambda(x_1, x_2, \ldots) M(dx).$$

down kernel

The measures $\{M_n\}$ are compatible with each other, through a certain canonical *down transition kernel* $p^{\downarrow}_{n+1,n}$ from $\mathbb{P}_{n+1}$ to $\mathbb{P}_n$.

down kernel

The measures $\{M_n\}$ are compatible with each other, through a certain canonical *down transition kernel* $p^{\downarrow}_{n+1,n}$ from $\mathbb{P}_{n+1}$ to $\mathbb{P}_n$.

$$p^{\downarrow}_{n+1,n}(\lambda, \mu) := \frac{\binom{n}{\mu_1,\ldots,\mu_{\ell(\mu)}}}{\binom{n+1}{\lambda_1,\ldots,\lambda_{\ell(\lambda)}}}, \text{ if } \lambda = \mu + \square$$

$$p^{\downarrow}_{n+1,n}(\lambda, \mu) := 0, \text{ otherwise}$$

down kernel

The measures $\{M_n\}$ are compatible with each other, through a certain canonical *down transition kernel* $p^{\downarrow}_{n+1,n}$ from $\mathbb{P}_{n+1}$ to $\mathbb{P}_n$.

$$p^{\downarrow}_{n+1,n}(\lambda,\mu) := \frac{\binom{n}{\mu_1,\dots,\mu_{\ell(\mu)}}}{\binom{n+1}{\lambda_1,\dots,\lambda_{\ell(\lambda)}}}, \text{ if } \lambda = \mu + \square$$

$$p^{\downarrow}_{n+1,n}(\lambda,\mu) := 0, \text{ otherwise}$$

$\binom{n}{\lambda_1,\lambda_2,\dots,\lambda_\ell}$ — number of paths in the Kingman graph from $\varnothing$ to $\lambda \in \mathbb{P}_n$.

down kernel

The measures $\{M_n\}$ are compatible with each other, through a certain canonical *down transition kernel* $p^{\downarrow}_{n+1,n}$ from $\mathbb{P}_{n+1}$ to $\mathbb{P}_n$.

$$p^{\downarrow}_{n+1,n}(\lambda,\mu) := \frac{\binom{n}{\mu_1,\dots,\mu_{\ell(\mu)}}}{\binom{n+1}{\lambda_1,\dots,\lambda_{\ell(\lambda)}}}, \text{ if } \lambda = \mu + \square$$

$$p^{\downarrow}_{n+1,n}(\lambda,\mu) := 0, \text{ otherwise}$$

$\binom{n}{\lambda_1,\lambda_2,\dots,\lambda_\ell}$ — number of paths in the Kingman graph from $\varnothing$ to $\lambda \in \mathbb{P}_n$.

$p^{\downarrow}_{n+1,n}(\lambda,\cdot)$ = take (uniformly) a random box from λ and delete it.

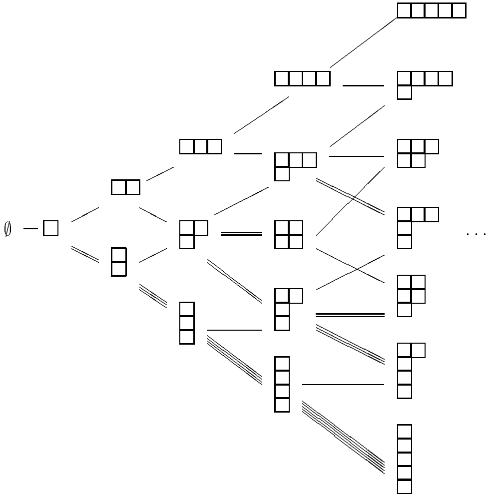

Figure: Kingman graph

Definition

$\{M_n\}$ — sequence of measures on $\mathbb{P}_n$ is called a *partition structure* if it is compatible with $p^{\downarrow}_{n+1,n}$:

$$M_{n+1} \circ p^{\downarrow}_{n+1,n} = M_n$$

Definition

$\{M_n\}$ — sequence of measures on $\mathbb{P}_n$ is called a *partition structure* if it is compatible with $p^{\downarrow}_{n+1,n}$:

$$M_{n+1} \circ p^{\downarrow}_{n+1,n} = M_n$$

Kingman's theorem = classification of partition structures

Definition

$\{M_n\}$ — sequence of measures on $\mathbb{P}_n$ is called a *partition structure* if it is compatible with $p^{\downarrow}_{n+1,n}$:

$$M_{n+1} \circ p^{\downarrow}_{n+1,n} = M_n$$

Kingman's theorem = classification of partition structures

- Any partition structure ⇒ up/down Markov chains.

Definition

$\{M_n\}$ — sequence of measures on $\mathbb{P}_n$ is called a *partition structure* if it is compatible with $p^{\downarrow}_{n+1,n}$:

$$M_{n+1} \circ p^{\downarrow}_{n+1,n} = M_n$$

Kingman's theorem = classification of partition structures

- Any partition structure $\Rightarrow$ up/down Markov chains.
- Which partition structures are good for obtaining diffusions (diffusions will be infinite-dimensional)?

1. de Finetti's Theorem
2. up/down Markov chains and limiting diffusions
3. Kingman's exchangeable random partitions
4. up/down Markov chains on partitions and limiting infinite-dimensional diffusions

Poisson-Dirichlet distributions

There is a distinguished two-parameter family of measures $PD(\alpha, \theta)$ on $\overline{\nabla}_\infty$, $\quad 0 \le \alpha < 1, \quad \theta > -\alpha$.

Poisson-Dirichlet distributions

There is a distinguished two-parameter family of measures $PD(\alpha, \theta)$ on $\overline{\nabla}_\infty$, $\quad 0 \le \alpha < 1, \quad \theta > -\alpha$.

- $\alpha = 0$ — introduced by Kingman, definition through a Poisson point process on $(0, \infty)$.

Poisson-Dirichlet distributions

There is a distinguished two-parameter family of measures $PD(\alpha, \theta)$ on $\overline{\nabla}_\infty$, $\quad 0 \leq \alpha < 1, \quad \theta > -\alpha$.

- $\alpha = 0$ — introduced by Kingman, definition through a Poisson point process on $(0, \infty)$.
- $\alpha \neq 0$ — Pitman(-Yor), '92–'95, motivated by stochastic processes. Definition through a Cox point process on $(0, \infty)$.

Poisson-Dirichlet distributions

There is a distinguished two-parameter family of measures $PD(\alpha, \theta)$ on $\overline{\nabla}_\infty$, $\quad 0 \le \alpha < 1, \quad \theta > -\alpha$.

- $\alpha = 0$ — introduced by Kingman, definition through a Poisson point process on $(0, \infty)$.
- $\alpha \neq 0$ — Pitman(-Yor), '92–'95, motivated by stochastic processes. Definition through a Cox point process on $(0, \infty)$.

Ewens-Pitman sampling formula

Partition structure corresponding to $PD(\alpha, \theta)$:

$$M_n(\lambda) = \frac{n!}{(\theta)_n} \cdot \frac{\theta(\theta+\alpha)\dots(\theta+(\ell(\lambda)-1)\alpha)}{\prod \lambda_i! \prod [\lambda : k]!} \cdot \prod_{i=1}^{\ell(\lambda)} \prod_{j=2}^{\lambda_i} (j-1-\alpha)$$

- Having a partition structure $\{M_n\}$ corresponding to $PD(\alpha, \theta)$, define up/down Markov chains $T_n^{(\alpha,\theta)}$ on $\mathbb{P}_n$ as before. They are reversible with respect to M_n.

- Having a partition structure $\{M_n\}$ corresponding to $PD(\alpha, \theta)$, define up/down Markov chains $T_n^{(\alpha,\theta)}$ on $\mathbb{P}_n$ as before. They are reversible with respect to M_n.
- One step of $T_n^{(\alpha,\theta)}$ = move one box in the corresponding Young diagram from one place to another.

- Having a partition structure $\{M_n\}$ corresponding to $PD(\alpha, \theta)$, define up/down Markov chains $T_n^{(\alpha,\theta)}$ on $\mathbb{P}_n$ as before. They are reversible with respect to M_n.
- One step of $T_n^{(\alpha,\theta)}$ = move one box in the corresponding Young diagram from one place to another.
- *Remark:* For $\alpha = 0$, these chains have population-genetic interpretation, and the limiting infinite-dimensional diffusions were constructed by Ethier and Kurtz in '81 (by approximating by finite-dimensional diffusions).

up-down Markov chains

- Having a partition structure $\{M_n\}$ corresponding to $PD(\alpha,\theta)$, define up/down Markov chains $T_n^{(\alpha,\theta)}$ on $\mathbb{P}_n$ as before. They are reversible with respect to M_n.
- One step of $T_n^{(\alpha,\theta)}$ = move one box in the corresponding Young diagram from one place to another.
- *Remark:* For $\alpha = 0$, these chains have population-genetic interpretation, and the limiting infinite-dimensional diffusions were constructed by Ethier and Kurtz in '81 (by approximating by finite-dimensional diffusions).
- Scale space by $1/n$ (embed $\mathbb{P}_n$ into $\overline{\nabla}_\infty$) and scale time by $1/n^2$.

up-down Markov chains

- Having a partition structure $\{M_n\}$ corresponding to $PD(\alpha, \theta)$, define up/down Markov chains $T_n^{(\alpha,\theta)}$ on $\mathbb{P}_n$ as before. They are reversible with respect to M_n.
- One step of $T_n^{(\alpha,\theta)}$ = move one box in the corresponding Young diagram from one place to another.
- *Remark:* For $\alpha = 0$, these chains have population-genetic interpretation, and the limiting infinite-dimensional diffusions were constructed by Ethier and Kurtz in '81 (by approximating by finite-dimensional diffusions).
- Scale space by $1/n$ (embed $\mathbb{P}_n$ into $\overline{\nabla}_\infty$) and scale time by $1/n^2$.
  The measures M_n converge to $PD(\alpha, \theta)$, what about Markov chains $T_n^{(\alpha,\theta)}$?

limiting infinite-dimensional diffusions

Theorem [P.]

1. As $n \to +\infty$, under the space and time scalings, the Markov chains $T_n^{(\alpha,\theta)}$ converge to an *infinite-dimensional diffusion process* $(X_{\alpha,\theta}(t))_{t \geq 0}$ on $\overline{\nabla}_\infty$.

limiting infinite-dimensional diffusions

Theorem [P.]

1. As $n \to +\infty$, under the space and time scalings, the Markov chains $T_n^{(\alpha,\theta)}$ converge to an *infinite-dimensional diffusion process* $(X_{\alpha,\theta}(t))_{t\geq 0}$ on $\overline{\nabla}_\infty$.
2. The Poisson-Dirichlet distribution $PD(\alpha,\theta)$ is the *unique invariant probability distribution* for $X_{\alpha,\theta}(t)$. The process is *reversible* and *ergodic* with respect to $PD(\alpha,\theta)$.

limiting infinite-dimensional diffusions

Theorem [P.]

1. As $n \to +\infty$, under the space and time scalings, the Markov chains $T_n^{(\alpha,\theta)}$ converge to an *infinite-dimensional diffusion process* $(X_{\alpha,\theta}(t))_{t\geq 0}$ on $\overline{\nabla}_\infty$.
2. The Poisson-Dirichlet distribution $PD(\alpha,\theta)$ is the *unique invariant probability distribution* for $X_{\alpha,\theta}(t)$. The process is *reversible* and *ergodic* with respect to $PD(\alpha,\theta)$.
3. The *generator* of $X_{\alpha,\theta}$ is explicitly computed:

$$\sum_{i,j=1}^{\infty} x_i(\delta_{ij} - x_j)\frac{\partial^2}{\partial x_i \partial x_j} - \sum_{i=1}^{\infty}(\theta x_i + \alpha)\frac{\partial}{\partial x_i}.$$

It acts on *continuous symmetric polynomials* in the coordinates $x_1, x_2, \ldots$.

limiting infinite-dimensional diffusions

Theorem [P.]

4. The spectrum of the generator in $L^2(\overline{\nabla}_\infty, PD(\alpha,\theta))$ is
$$\{0\} \cup \{-n(n-1+\theta)\colon n = 2, 3, \ldots\},$$
the eigenvalue 0 is simple, and the multiplicity of $-n(n-1+\theta)$ is the number of partitions of n with all parts ≥ 2.

limiting infinite-dimensional diffusions

Theorem [P.]

4. The spectrum of the generator in $L^2(\overline{\nabla}_\infty, PD(\alpha,\theta))$ is
$$\{0\} \cup \{-n(n-1+\theta)\colon n = 2, 3, \dots\},$$
the eigenvalue 0 is simple, and the multiplicity of $-n(n-1+\theta)$ is the number of partitions of n with all parts ≥ 2.

Remark: connection with the Pascal triangle model

Degenerate parameters

$$\alpha < 0 \text{ arbitrary}, \qquad \theta = -2\alpha$$

$\Rightarrow$ partitions have ≤ 2 parts.
de Finetti's model with (any) symmetric Beta distribution.

limiting infinite-dimensional diffusions

Theorem [P.]

4. The spectrum of the generator in $L^2(\overline{\nabla}_\infty, PD(\alpha, \theta))$ is
$$\{0\} \cup \{-n(n-1+\theta) \colon n = 2, 3, \dots\},$$
the eigenvalue 0 is simple, and the multiplicity of $-n(n-1+\theta)$ is the number of partitions of n with all parts ≥ 2.

Remark: connection with the Pascal triangle model

Degenerate parameters

$$\alpha < 0 \text{ arbitrary}, \qquad \theta = -2\alpha$$

$\Rightarrow$ partitions have ≤ 2 parts.
de Finetti's model with (any) symmetric Beta distribution.

$\alpha < 0$, $\theta = -K\alpha \Rightarrow K$-dimensional generalization.

No finite-dimensional approximating diffusions for $\alpha \neq 0$!

No finite-dimensional approximating diffusions for $\alpha \neq 0$!

1. The transition operators of the Markov chains $T_n^{(\alpha,\theta)}$ act on symmetric functions in the coordinates $\lambda_1, \ldots, \lambda_\ell$ of a partition $\lambda \in \mathbb{P}_n$.

scheme of proof

No finite-dimensional approximating diffusions for $\alpha \neq 0$!

1. The transition operators of the Markov chains $T_n^{(\alpha,\theta)}$ act on symmetric functions in the coordinates $\lambda_1, \ldots, \lambda_\ell$ of a partition $\lambda \in \mathbb{P}_n$.
2. Write the operators $T_n^{(\alpha,\theta)}$ in a suitable basis (closely related to the monomial symmetric functions).

scheme of proof

No finite-dimensional approximating diffusions for $\alpha \neq 0$!

1. The transition operators of the Markov chains $T_n^{(\alpha,\theta)}$ act on symmetric functions in the coordinates $\lambda_1, \ldots, \lambda_\ell$ of a partition $\lambda \in \mathbb{P}_n$.
2. Write the operators $T_n^{(\alpha,\theta)}$ in a suitable basis (closely related to the monomial symmetric functions).
3. Pass to $n \to +\infty$ limit of generators (*algebraically!*). The processes' core is the algebra of symmetric functions.

scheme of proof

No finite-dimensional approximating diffusions for $\alpha \neq 0$!

1. The transition operators of the Markov chains $T_n^{(\alpha,\theta)}$ act on symmetric functions in the coordinates $\lambda_1, \ldots, \lambda_\ell$ of a partition $\lambda \in \mathbb{P}_n$.
2. Write the operators $T_n^{(\alpha,\theta)}$ in a suitable basis (closely related to the monomial symmetric functions).
3. Pass to $n \to +\infty$ limit of generators (*algebraically!*). The processes' core is the algebra of symmetric functions.
4. Use general technique of Trotter-Kurtz to deduce convergence of the processes

Thank you for your attention

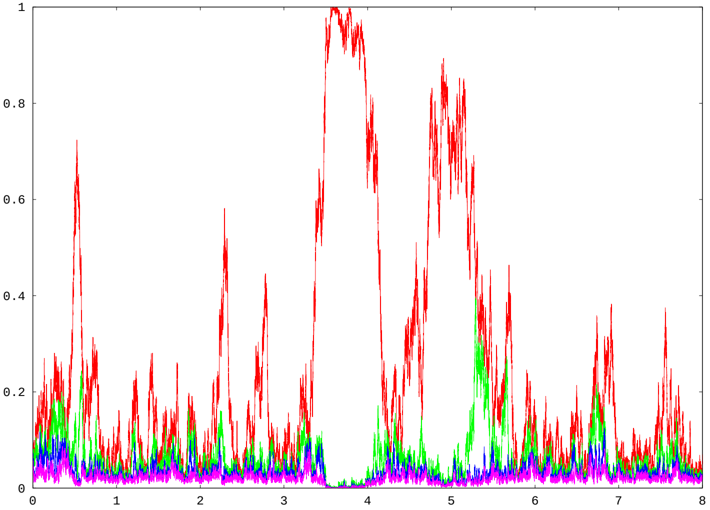

Figure: $x_1(t) \geq x_2(t) \geq x_3(t) \geq x_4(t)$