Probability at Warwick Young Researchers Workshop

Trickle-down growth models, Doob-Martin boundaries, and random matrices

Steven N. Evans

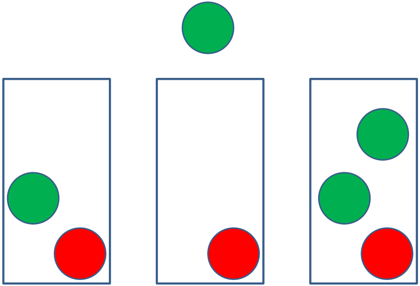
University of California at Berkeley

July, 2010

THIS VERSION OF THE NOTES IS INCOMPLETE.
I WILL ADD TO THEM DURING THE WORKSHOP.

Classical Pólya urn

- Start with a red and green ball in an urn.
- At each point in time, pick a ball uniformly at random from the urn and replace it along with one of the same color.



- Let μ be a finite measure on a Polish space E.
- Say that a sequence $\{X_n\}_{n=1}^{\infty}$ of r.v. with values in E is a Pólya sequence with parameter μ if for every Borel set $B \subseteq E$

$$\mathbb{P}\{X_1 \in B\} = \frac{\mu(B)}{\mu(E)} \tag{1}$$

and

$$\mathbb{P}\{X_{n+1} \in B \,|\, X_1, \ldots, X_n\} = \frac{\mu_n(B)}{\mu_n(E)} \tag{2}$$

where

$$\mu_n := \mu + \sum_{i=1}^{n} \delta_{X_i}$$

and δ_x denotes the unit point mass at x.

- The successive draws in the classical Pólya urn is a Pólya sequence with $E = \{\mathrm{red}, \mathrm{green}\}$ and $\mu\{\mathrm{red}\} = \mu\{\mathrm{green}\} = 1$.

- For finite E, the sequence $\{X_n\}_{n=1}^\infty$ represents the results of successive draws from an urn where initially the urn has "$\mu\{x\}$ balls of color x" and, after each draw, the ball drawn is replaced and another ball of its same color is added to the urn.
- Without the restriction to finite E, for any Borel measurable function ϕ from E to another space F, the sequence $\{\phi(X_n)\}_{n=1}^\infty$ is a Pólya sequence with parameter $\mu \circ \phi^{-1}$, where $\mu \circ \phi^{-1}(A) = \mu\{x \in E : \phi(x) \in A\}$.

- Recall that a random variable $(Z_1, \ldots, Z_n)$ taking values in the simplex

$$\{(z_1, \ldots, z_n) : z_k \geq 0 \ \forall k, \ \sum_k z_k = 1\}$$

has a Dirichlet distribution with parameter $(\alpha_1, \ldots, \alpha_n)$ if $(Z_1, \ldots, Z_{n-1})$ has density

$$\frac{\Gamma(\sum_k \alpha_k)}{\prod_k \Gamma(\alpha_k)} \prod_k z_k^{\alpha_k - 1},$$

where $z_n := (1 - \sum_{k=1}^{n-1} z_k)$.

- Exercise: Compute the joint moment $\mathbb{P}[\prod_k Z_k^{m_k}]$ for $m_k \in \mathbb{N}_0$.
Note: I will usually use $\mathbb{P}[\cdot]$ for expected value.

Ferguson random measures

- Let μ be a finite Borel measure on the Polish space E.
- A random probability measure μ^* on E has a Ferguson distribution with parameter μ if for every finite partition $(B_1, \ldots, B_r)$ of E the vector $(\mu^*(B_1), \ldots, \mu^*(B_r))$ has a Dirichlet distribution with parameter $(\mu(B_1), \ldots, \mu(B_r))$ (when $\mu(B_i) = 0$, this means $\mu^*(B_i) = 0$ with probability 1).

Blackwell – MacQueen theorem

Theorem 1

Let $\{X_n\}_{n=1}^{\infty}$ *be a Pólya sequence with parameter* μ*. Then,*

(i) $\mu_n/\mu_n(E)$ *converges with probability* 1 *as* $n \to \infty$ *to a limiting discrete measure* μ^**,*

(ii) μ^* *has a Ferguson distribution with parameter* μ*,*

(iii) *given* μ^**, the variables* $X_1, X_2, \ldots$ *are independent with distribution* μ^**.*

Proof. Suppose first that E is finite, say $E = \{1, 2, ..., r\}$. Let μ^*, $\{X_n\}_{n=1}^\infty$ be variables whose joint distribution is defined by (ii) and (iii). If π_n is the empirical distribution of $X_1, \ldots, X_n$, then it follows from the strong law of large numbers that $\pi_n \to \mu^*$ with probability 1 as $n \to \infty$. Since

$$\frac{\mu_n}{\mu_n(E)} = \frac{\mu + n\pi_n}{\mu(E) + n},$$

(i) follows.

It remains to show that $\{X_n\}_{n=1}^{\infty}$ is a Pólya sequence with parameter μ, which is equivalent to

$$\mathbb{P}(A) = \prod_x \frac{\mu(x)^{[n(x)]}}{\mu(E)^{[n]}}, \tag{3}$$

where $A := \{X_1 = x_1, \ldots, X_n = x_n\}$, $n(x)$ denotes the number of i with $x_i = x$, and $a^{[k]} = a(a+1)\ldots(a+k-1)$. Since

$$\mathbb{P}(A \,|\, \mu^*) = \prod_x \mu^*(x)^{n(x)},$$

we get

$$\mathbb{P}(A) = \mathbb{P}\left[\prod_x \mu^*(x)^{n(x)}\right]. \tag{4}$$

That the right sides of (3) and (4) are equal is just the formula for the moments of Dirichlet distributions. The case of general E follows by a straightforward approximation procedure. □

One perspective on the Blackwell – MacQueen theorem

- Recall that when E is finite,

$$\mathbb{P}\{X_1 = x_1, \ldots, X_n = x_n\} = \prod_x \frac{\mu(x)^{[n(x)]}}{\mu(E)^{[n]}},$$

  where $n(x) := \{1 \leq i \leq n : x_i = x\}$. Note the symmetry. The symmetry holds for general E.

- If $\mathbb{Q}$ is the distribution of $\{X_n\}_{n=1}^{\infty}$ and $\mathbb{F}^\mu$ is the distribution of a Ferguson distribution with parameter μ, then

$$\mathbb{Q}(B) = \int \nu^{\otimes\infty}(B)\, \mathbb{F}^\mu(d\nu).$$

  That is, $\mathbb{Q}$ is a convex combination of product measures with identical factors.

Exchangeability and spreadability

- Consider an infinite random sequence $\xi = \{\xi_n\}_{n=1}^{\infty} \in E^{\infty}$ where E is a Polish space.
- Say that ξ is exchangeable if

$$(\xi_{k_1}, \xi_{k_2}, \dots) \stackrel{d}{=} (\xi_1, \xi_2, \dots) \qquad (5)$$

  for any finite permutation $(k_1, k_2, \dots)$ of $\mathbb{N}$.
- Say that ξ is spreadable if (5) holds for all strictly increasing sequences $k_1 < k_2 < \dots$
- Note: exchangeable $\Rightarrow$ spreadable $\Rightarrow$ stationary.

- If ξ is stationary, an event $A \in \mathcal{F}$ is invariant if $A = \{\omega \in \Omega : (\xi_1(\omega), \xi_2(\omega), \ldots) \in B\}$ for some Borel set $B \subseteq E^\infty$ such that
$$\mathbb{P}\left(\{(\xi_1, \xi_2, \ldots) \in B\} \triangle \{(\xi_2, \xi_3, \ldots) \in B\}\right) = 0.$$
- Exercise: Show that the invariant events form a σ-field $\mathcal{I}_\xi$ and that this σ-field is generated by the limits
$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} f(\xi_k, \ldots, \xi_{k+m-1})$$
for $m \in \mathbb{N}$ and bounded Borel functions $f : E^m \to \mathbb{R}$. (Why do these limits exist?)

Theorem 2

The following conditions are equivalent for any infinite random sequence ξ *taking values in a Polish space* E*:*

(i) ξ *is exchangeable;*

(ii) ξ *is spreadable;*

(iii) $P\{\xi \in \cdot \mid \eta\} = \eta \otimes \eta \otimes \cdots$ *a.s. for some random probability measure* η *on* E.

The random measure η *is then a.s. unique and equals* $P\{\xi_1 \in \cdot \mid \mathcal{I}_\xi\}$.

Exchangeable probability measures are convex combinations of product measures with identical factors.

Convex sets and extreme points

- Suppose that E is a real vector space and $A \subseteq E$.
- A point x of A is an extreme point of A if the relations
  - $a \in A$, $b \in A$,
  - $x = (1 - \lambda)a + \lambda b$,
  - $0 \leq \lambda \leq 1$,

  together entail that x is either a or b.
- Write $\mathrm{ex}A$ for the extreme points of A.

Exercise

Show that the extreme points of the convex set of exchangeable probability measures on E^∞, where E is a Polish space, are the product measures with identical factors.

Theorem 3

Let E be a real, Hausdorff, locally convex, topological vector space and K a nonempty, compact, convex subset of E. Then K is the closed convex hull in E of the set of extreme points of K.

Note: A consequence of the Krein-Milman theorem is that any point $x \in K$ can be represented as

$$x = \int y \, \lambda(dy)$$

$$\Leftrightarrow$$

$$\phi(x) = \int_{\mathrm{ex}K} \phi(y) \, \lambda(dy), \quad \forall \phi \in E^*$$

for some probability measure λ supported on the closure of $\mathrm{ex}K$.

- Suppose that E is the Hilbert space $\ell^2 := \{(x_1, x_2, \ldots) \in \mathbb{R}^{\mathbb{N}} : \sum_n x_n^2 < \infty\}$. Put $K := \{(x_1, x_2, \ldots) \in E : \sum_n 2^n x_n^2 \leq 1\}$. Show that K is compact and convex. Show that $\mathrm{ex}K = \{(x_1, x_2, \ldots) \in E : \sum_n 2^n x_n^2 = 1\}$ and that the closure of $\mathrm{ex}K$ is K, so Krein-Milman doesn't say much in this case.
- Give an example of a compact, convex subset $K \subset \mathbb{R}^d$ for which $\mathrm{ex}K$ is not closed.

Exercise – Birkhoff's theorem

- Suppose that P is an $n \times n$ matrix that is doubly stochastic; that is, P has non-negative entries and each row and column add to 1. Use Krein-Milman to show that there are permutation matrices Π_k and $p_k \geq 0$ with $\sum_k p_k = 1$ such that $P = \sum_k p_k \Pi_k$ (a permutation matrix is a matrix that has a single 1 in each row and column and zeros elsewhere). Hint: Note that if P is not a permutation matrix, then for some N there are pairs of indices $(i_1, j_1), (i_2, j_2), \ldots, (i_{2N}, j_{2N})$ such that $P(i_k, j_k) > 0$, $i_k = i_{k+1}$ if k is odd, and $j_k = j_{k+1}$ if k is even (with the convention $2N + 1 = 1$).
- Recall that $[n] := \{1, 2, \ldots, n\}$. Set

$$S := \{(k_1, \ldots, k_n) \in [n]^n : k_i \neq k_j,\, i \neq j\}.$$

Show that there is an S-valued Markov chain $\{X_m\}_{m=0}^{\infty} = \{(X_m(1), \ldots, X_m(n))\}_{m=0}^{\infty}$ such that for each i the process $\{X_m(i)\}_{m=0}^{\infty}$ is a Markov chain with transition matrix P.

Uniqueness?

- de Finetti's theorem asserts that an exchangeable probability measure is a unique convex combination of product measures.
- If K is a convex subset of $\mathbb{R}^d$ such that each $x \in K$ is a unique convex combination of points in $\mathrm{ex}K$, then K is a simplex.
- How is the set of exchangeable probability measures like a simplex?

Another perspective on the Blackwell – MacQueen theorem

- Let $\{X_n\}_{n=1}^{\infty}$ be a Pólya sequence with parameter μ.
- Put $Y_n := \sum_{k=1}^{n} \delta_{X_k}$ and $Y_0 := 0$. Show that $\{Y_n\}_{n=0}^{\infty}$ is a Markov chain (Exercise). Write $P(y, dz) := \mathbb{P}\{Y_{n+1} \in dz \mid Y_n = y\}$.
- Recall that $\mu^* := \lim_{n\to\infty} \frac{1}{n} Y_n$ exists and the invariant σ-field of $\{X_n\}_{n=1}^{\infty}$ = tail σ-field of $\{Y_n\}_{n=0}^{\infty} = \sigma(\mu^*) =: \mathcal{T}$.
- If $Z = \Phi(\mu^*)$ is a bounded, non-negative, $\mathcal{T}$-measurable r.v., then

$$\mathbb{P}[Z \mid Y_0, \ldots, Y_n] = \int \Phi(\nu)\, \mathbb{F}^{\mu+Y_n}(d\nu) =: \Psi(Y_n)$$

  is a martingale and

$$\int \Psi(z) P(y, dz) = \Psi(y)$$

  – the function Ψ is regular or harmonic.
- Conversely, if Ψ is a bounded, non-negative, harmonic function, then $\{\Psi(Y_n)\}_{n=0}^{\infty}$ is a martingale, $\Psi(Y_n) = \mathbb{P}[Z \mid Y_0, \ldots, Y_n]$ for some $\mathcal{T}$-measurable Z, and $Z = \Phi(\mu^*)$ for some Φ.

- Recall that
$$y \mapsto \int \Phi(\nu)\, \mathbb{F}^{\mu+y}(d\nu)$$
is a harmonic function.
- Note that
$$\int \Phi(\nu)\, \mathbb{F}^{\mu+y}(d\nu) = \int \Phi(\nu)\, \frac{d\mathbb{F}^{\mu+y}}{d\mathbb{F}^{\mu}}(\nu)\, \mathbb{F}^{\mu}(d\nu).$$
- This suggests that
$$y \mapsto \frac{d\mathbb{F}^{\mu+y}}{d\mathbb{F}^{\mu}}(\nu)$$
is a non-negative harmonic function for each ν and any non-negative harmonic functions is a unique non-negative "linear combination" of these functions.
- Equivalently, the non-negative harmonic functions that take the value 1 at the measure 0 form a convex set, and perhaps an arbitrary such function h has a representation as
$$y \mapsto \int \frac{d\mathbb{F}^{\mu+y}}{d\mathbb{F}^{\mu}}(\nu)\, \pi(d\nu)$$
for some unique probability measure π on the set of probability measures on E.

Exercise

- Suppose that $E = \{1, 2, \dots, r\}$, so that the state space of $Y_n = \sum_{k=1}^n \delta_{X_k}$ may be thought of as $(\mathbb{N}_0)^r$.
- Show that

$$h_\nu(y) := \frac{d\mathbb{F}^{\mu+y}}{d\mathbb{F}^\mu}(\nu) = \frac{(\sum_k(\mu_k + y_k) - 1)_{\sum_k y_k}}{\prod_k(\mu_k + y_k - 1)_{y_k}} \prod_k \nu_k^{y_k},$$

where $(a)_\ell := a(a-1)\cdots(a-\ell+1)$ is the usual Pochhammer symbol.

- Show by direct calculation that the function h_ν is harmonic for Y.
- Let $P(y, z)$ be the transition matrix for Y. Show that

$$\frac{1}{h_\nu(y)} P(y, z) h_\nu(z)$$

is also a transition matrix. What is the corresponding Markov chain?

Choquet Theory – existence

- Suppose that E is a locally convex topological vector space and K is a non-empty, metrizable, compact, convex subset of E.
- Let μ be a probability measure on K. A point x in E is said to be a barycenter of μ if $f(x) = \int_K f \, d\mu$ for every continuous linear functional f on E. (We will sometimes write $\mu(f)$ for $\int_K f d\mu$.)
- Fact: Each μ has a unique barycenter. (Exercise)
- If μ is a measure on K and S is a Borel subset of K, we say that μ is supported by S if $\mu(K \backslash S) = 0$.

Proposition 4

The extreme points of the compact, convex set K form a G_δ set.

Proof: Suppose that the topology of K is given by the metric d, and for each integer $n \geq 1$ let $F_n := \{x : x = 2^{-1}(y+z), y, z \in K, d(z,y) \geq 1/n\}$. It is easily checked that each F_n is closed, and that a point x of K is not extreme if and only if it is in some F_n. Thus, the complement of the extreme points is an F_σ. □

Proposition 5

Suppose that $x \in K$. Then x is an extreme point of K if and only if the point mass δ_x is the only probability measure on K with barycenter x.

Proof. Exercise.

Proposition 6

The compact, convex set K is affinely homeomorphic to a (norm-)compact, convex subset of a Hilbert space.

Proof. This is well-known. Try it as an exercise, but it is rather difficult.

Theorem 7

Each $x \in K$ is the barycenter of a probability measure μ supported by $\mathrm{ex}K$.

Proof. We may suppose that K is a compact convex subset of a Hilbert space $(H, \|\|)$. Put

$$\gamma := \sup \left\{ \int_K \|h\|^2 \, \mu(dh) : x = \int_K h \, \mu(dh) \right\}.$$

Find $\{\mu_n\}$ such that $x = \int_K h \, \mu_n(dh)$ for all n and $\gamma = \lim_n \int_K \|h\|^2 \, \mu(dh)$.
The probability measures on K are weak* compact, so we can suppose that $\mu_n \to \mu$ weak* for some μ.
Note that $x = \int_K h \, \mu(dh)$ and $\gamma = \int_K \|h\|^2 \, \mu(dh)$.

Consider $u \in K$ of the form $u = (v + w)/2$ for $v \neq w$. Then,

$$\frac{\|v\|^2 + \|w\|^2}{2} = \left\|\frac{v+w}{2}\right\|^2 + \left\|\frac{v-w}{2}\right\|^2 = \|u\|^2 + \left\|\frac{v-w}{2}\right\|^2.$$

Suppose that $y = \int_K h \, \sigma(dh)$ where $\sigma = \sum_k p_k \delta_{u_k}$ with $u_1 = u$. Put

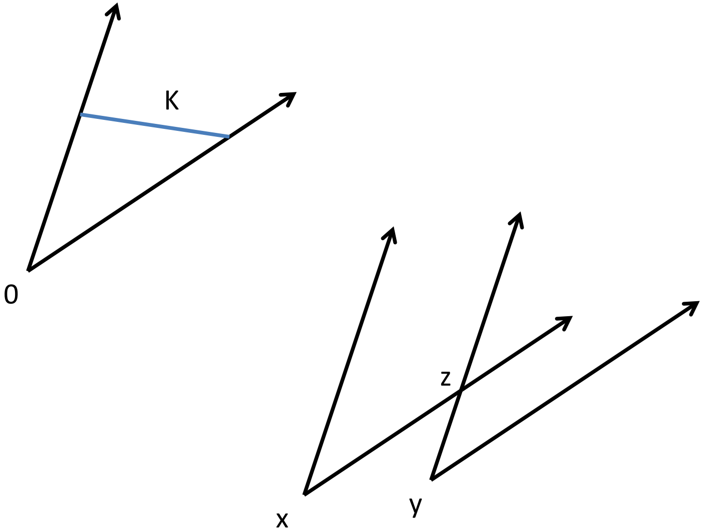
$$\tau := p_1 \frac{1}{2} (\delta_v + \delta_w) + \sum_{k>1} p_k \delta_{u_k}.$$

Then, $y = \int_K h \, \sigma(dh)$ and

$$\int_K \|h\|^2 \, \sigma(dh) < \int_K \|h\|^2 \, \tau(dh).$$

Consequently, if μ is not supported on $\mathrm{ex}K$ it can be "improved" (this takes some work involving approximation by discrete measures). □

- Assume wlog that K is contained in a closed hyperplane which misses the origin.
- The convex set $\tilde{K} := \{\alpha x : \alpha \geq 0, \, x \in K\}$ is the cone generated by K.
- The convex cone $\tilde{K}$ induces a translation invariant partial ordering on E: $x \geq y$ if and only if $x - y \in \tilde{K}$.
- We call K a simplex if $\tilde{K}$ is a lattice in the partial order induced by K; that is, if each pair x, y in $\tilde{K}$ has a least upper bound. (We say that z is the least upper bound for x and y if $z \leq w$ whenever $w \geq x$ and $w \geq y$.)


K
0
z
x
y

Exercise

Let $K := \{(x, y, z) \in \mathbb{R}^3 : |x| \leq 1, \ |y| \leq 1, \ z = 1\}$.

- Show that there are points of K that are not unique convex combinations of the extreme points.
- Show that $\tilde{K}$ is not a lattice.
- Show for there exists $a', a'' \geq 0$ and $b', b'' \in \mathbb{R}^3$ such that $(a'K + b') \cap (a''K + b'')$ is non-empty but not of the form $aK + b$ for some $a \geq 0$ and $b \in \mathbb{R}^3$.

Theorem 8

Suppose that K is a non-empty, metrizable, compact, convex subset of a locally convex space E. The following assertions are equivalent

(1) *K is a simplex.*

(2) *For all $a', a'' \geq 0$ and $b', b'' \in E$, either $(a'K + b') \cap (a''K + b'')$ is empty or of the form $aK + b$ for some $a \geq 0$ and $b \in E$.*

(3) *For each $x \in K$ there is a unique probability measure μ supported on $\mathrm{ex}K$ such that μ has barycenter x.*

Let I be a countable set and let Π be a sub-stochastic $I \times I$ matrix. Define the Green kernel Γ of Π as the $I \times I$ matrix

$$\Gamma(i,j) := \sum_{n=0}^{\infty} \Pi^n(i,j) \le \infty$$

so that, formally, $\Gamma = (I - \Pi)^{-1}$.

Markov chains

Let $X = \{X_n\}_{n=0}^{\infty}$ be a Markov chain on I (with coffin state ∂ adjoined) with 1-step transition matrix Π. That is,

$$\mathbb{P}^{i_0}\{X_1 = i_1, \ldots, X_n = i_n\} = \Pi(i_0, i_1)\Pi(i_1, i_2)\ldots\Pi(i_{n-1}, i_n).$$

Then,

$$\Gamma(i, j) = \mathbb{P}^i[\text{time spent by } X \text{ in } j].$$

ASSUMPTION. There exists a reference point b in I such that

$$0 < \Gamma(b, j) < \infty, \ \forall \ j \in I.$$

Consequently,

(i) the chain can get to any state from state b;

(ii) every state is transient.

- The strong Markov property of X shows that

$$\Gamma(i,j) = \mathbb{P}^i\{D_j < \infty\}\Gamma(j,j) \le \Gamma(j,j),\ \forall\ i,j, \tag{6}$$

where

$$D_j := \inf\{n \ge 0 : X_n = j\}.$$

- Define the Martin kernel κ on $I \times I$ as

$$\kappa(i,j) := \frac{\Gamma(i,j)}{\Gamma(b,j)}. \tag{7}$$

- Exercise: Show that

$$\kappa(i,j) = \frac{\mathbb{P}^i\{X \text{ hits } j\}}{\mathbb{P}^b\{X \text{ hits } j\}}.$$

It follows from (6) (Exercise) that

$$\kappa(i,j) \le \kappa(j,j) < \infty, \ \forall\ i,j. \tag{8}$$

It is another easy consequence of the strong Markov property (Exercise) that

$$\kappa(i,j) \le \kappa(i,i) < \infty, \ \forall\ i,j. \tag{9}$$

A function f from I to $\mathbb{R}$ is called excessive (respectively, regular) (for Π) if

(i) $0 \leq f < \infty$;

(ii) $\Pi f \leq f$ (respectively $\Pi f = f$).

The set of excessive functions forms a cone C in $\mathbb{R}^I$. The topology of C is the one induced by that of $\mathbb{R}^I$, that is, the topology of pointwise convergence.

Exercise: Show that if f is excessive (resp. regular), then $\{f(X_n)\}_{n=0}^{\infty}$ is a super-martingale (resp. martingale).

Because of our standing assumption,

$$\theta(j) := \sup_n \Pi^n(b, j) > 0,$$

and since a function f in C satisfies

$$f \geq \Pi f \geq \Pi^2 f \geq \ldots,$$

we have

$$f(b) \geq \theta(j) f(j), \ \forall j. \tag{10}$$

In particular, every f in C may be written as

$$f = f(b) f^*, \quad f^* \in S := \{f \in C : f(b) = 1\}.$$

The study of the cone C thus reduces to the study of its section S.

Representation of excessive function – existence

Proposition 9

The set S is a compact convex metrizable subset of the locally convex linear topological space $\mathbb{R}^I$.

Theorem 10

If $f \in S$ then there exists a probability measure ν on supported on $\mathrm{ex}S$ such that

$$f(i) = \int_{\mathrm{ex}S} \xi(i)\,\nu(d\xi) \tag{11}$$

Theorem 11

Furthermore, the measure ν is uniquely determined by the excessive function f.

We know Theorem 11 will follow from the next lemma.

Lemma 12

The cone C is a lattice in its intrinsic order.

Note: The intrinsic order $\ll$ on C is defined as follows: for $x, z \in C$, we write $x \ll z$ if $\exists u \in C$ with $x + u = z$.

- Let μ be a (non-negative) measure on I such that

$$\Gamma\mu(i) := \sum_{j \in I} \Gamma(i,j)\mu(j) < \infty, \ \forall\, i.$$

Then $\Gamma\mu$ is called the potential (due to the charge μ).

- Since

$$\Pi\Gamma\mu = \Gamma\mu - \mu \leq \Gamma\mu, \tag{12}$$

the function $\Gamma\mu$ is excessive.

- The equation

$$\mu = \Gamma\mu - \Pi\Gamma\mu \tag{13}$$

determines μ from $\Gamma\mu$, and

$$\Pi^n\Gamma\mu = \sum_{k \geq n} \Pi^k\mu \downarrow 0 \text{ as } n \to \infty. \tag{14}$$

Riesz decomposition theorem

Theorem 13

If f is excessive, then f has a unique decomposition

$$f = \nu + \Gamma\mu \tag{15}$$

where ν is regular and μ is a measure on I. Indeed,

$$\nu = \lim_n \Pi^n f, \tag{16}$$

$$\mu = f - \Pi f \tag{17}$$

Proof: Define μ by (17). Then $\mu(i) \geq 0$, $\forall\ i$, and

$$(I + \Pi + \ldots + \Pi^n)\mu = f - \Pi^{n+1} f.$$

The Monotone Convergence Theorem yields (15) with ν as is (16). The properties (13) and/or (14) make the uniqueness assertion obvious. $\square$

We can now prove Lemma 12 by showing that if

$$f_1 = \nu_1 + \Gamma\mu_1, \quad f_2 = \nu_2 + \Gamma\mu_2,$$

then the greatest lower bound and least upper bound operations $\wedge\wedge$ and $\vee\vee$ witnessing the lattice structure of C in its intrinsic order is exhibited by the equations

$$f_1 \wedge \wedge f_2 = \lim_n \Pi^n(\nu_1 \wedge \nu_2) + \Gamma(\mu_1 \wedge \mu_2),$$

$$f_1 \vee \vee f_2 = f_1 + f_2 - f_1 \wedge \wedge f_2,$$

where

$$(\nu_1 \wedge \nu_2)(i) := \nu_1(i) \wedge \nu_2(i), \quad (\mu_1 \wedge \mu_2)(i) := \mu_1(i) \wedge \mu_2(i)$$

Exercise

Verify the above claim. It may help to note that any regular function dominated by a potential (in the usual partial order on functions) is zero and so if ν is a regular function with $\nu \leq f_1 \wedge f_2$, then $\nu \leq \lim_n \Pi^n(\nu_1 \wedge \nu_2)$.

Exercise

Suppose that $I = \mathbb{N} \times \mathbb{N}$, $\Pi((i,1),(i+1,1)) = \frac{1}{2}$, $\Pi((i,1),(i,2)) = \frac{1}{2}$, and $\Pi((i,j),(i,j+1)) = 1$ for $j \geq 2$.

- Compute Γ and κ for this chain (there is only one possible choice of reference point b).
- Describe the regular functions explicitly.
- Compute explicitly the greatest lower bound of two regular functions in the intrinsic order.

Proposition 14

For each j in I the function $\kappa(\cdot, j)$ is a (non-regular) extremal element of S. Every extremal element of S that is not of the form $\kappa(\cdot, j)$ for some j in I is regular.

- Since potential determines charge, the map

$$\phi : I \to S \subset \mathbb{R}^I, \quad \phi(j) := \kappa(\cdot, j)$$

  is injective.
- Identify I with $\phi(I)$ and let F be the compact closure of $I(=\phi(I))$ in S. The set F is called the Martin compactification of I.
- Since the topology of F is inherited from that of $\mathbb{R}^I$, the following holds: For each i, the map $\kappa(i, \cdot)$ extends continuously to F, so we have a map $\kappa : I \times F \to \mathbb{R}_+$.
- For $\xi \in F \backslash I$, we sometimes use the alternative notation $\kappa(i, \xi)$ for $\xi(i)$.

Doob-Hunt theorem

Theorem 15

Every extremal element of S is of the form $\kappa(\cdot, \xi)$ for some ξ in F. Let F_e be the set of ξ in F for which $\kappa(\cdot, \xi)$ is extremal. Then each f in S can be written uniquely as

$$f = \int_{F_e} \kappa(\cdot, \xi)\, \nu(d\xi) = \nu + \Gamma\mu$$

where ν is a probability measure on F_e,

$$\nu := \int_{F_e \setminus I} \kappa(\cdot, \xi)\, \nu(d\xi)$$

is regular, and

$$\mu(j) := [\Gamma(b, j)]^{-1} \nu(j).$$

- Once we establish the first sentence of Theorem 15, the remainder of the theorem follows from the Choquet results (10) and (11) and we then have $\mathrm{ex} S = F_e \subset F$.
- It is enough to prove every element f of S may be written as

$$f = \int_F \kappa(\cdot, \xi)\, \nu(d\xi)$$

for some (not necessarily unique) probability measure ν on F.
- This is because it will follow that S is the closed convex hull of F, and the following result shows that the extremal elements of S are contained in F.

Proposition 16

Suppose that E is a Hausdorff LCTVS and A is a compact subset of E whose closed convex hull K is compact. Then each extreme point of K belongs to A.

Proof. Let x be an extreme point of K. If U is any closed convex neighborhood of 0 in E, then there exist finitely many points a_i of A, $1 \leq i \leq n$, such that the sets $a_i + U$ cover A. Let K_i be the closed convex hull of $A \cap (a_i + U)$. Each K_i is compact. The convex hull of the union of the K_i, being compact, contained in K, and containing A, must be K itself. Hence, $x = \sum_{i=1}^n \lambda_i x_i$ with x_i in K_i, $\lambda_i \geq 0$, and $\sum_{i=1}^n \lambda_i = 1$. Since x is an extreme point of K, x must coincide with x_i for some i. Thus x belongs to $K_i \subset a_i + U$, and so x belongs to $A + U$. Since A is closed and U is arbitrary, it follows that x belongs to A, as claimed. □

We have now reduced the problem of proving Theorem 15 to that of proving:

Claim: Every element f of S may be written as

$$f = \int_F \kappa(\cdot, \xi)\, \nu(d\xi)$$

for some (not necessarily unique) probability measure ν on F.

Proof of the Claim. Fix f in S. Choose a measure β such that

$$0 < \Gamma\beta(i) < \infty, \forall\ i.$$

By (8), it is enough to choose β so that $\beta(j) > 0$, $\forall\ j$, and $\sum \Gamma(j,j)\beta(j) < \infty$. Let

$$f_n(i) := \min(f(i), n\Gamma\beta(i)).$$

Then f_n is excessive, and since f_n is dominated by the potential $n\Gamma\beta$, it follows from (14) and the Riesz theorem that f_n is a potential:

$$f_n(i) = \sum \Gamma(i,j)\mu_n(j) = \sum \kappa(i,j)\nu_n(j), \qquad (18)$$

where $\nu_n = \Gamma(b,j)\mu_n(j)$.

Since $f_n(b) = f(b) = 1$ for large n, and $\kappa(b, j) = 1$, $\forall\ j$, it follows that (for large n) ν_n is a probability measure on F with $\nu_n(I) = 1$.
Since F is compact metrizable, $\Pr(F)$ is compact metrizable in the weak topology. Let ν be a sub-sequential limit of $\{\nu_n\}$ in $\Pr(F)$. Then the statement follows from (18). □

The following analytical problem remains:

How can we determine the 'extremal' part F_e of F?

Example: Simple random walk

Let X be the simple random walk on $\mathbb{Z}^d$, $d \geq 3$, such that

$$\Pi(i,j) = \begin{cases} (2d)^{-1} & \text{if } |j-i| = 1 \\ 0 & \text{otherwise} \end{cases}$$

Our standing assumptions hold with the reference state $b = 0$. It is well known that

$$\Gamma(i,j) \sim \text{constant}|j-i|^{2-d} \quad (|j-i| \to \infty)$$

Since

$$\kappa(i,j) \sim |j-i|^{2-d}/|j|^{2-d}$$

it is clear that F is the one-point compactification $I \cup \{\infty\}$ of I and that $\kappa(i,\infty) = 1$, $\forall i$.

Thus, every regular function is constant.

Example: Space-time coin-tossing

Consider $X_n := (H_n, n)$ where H_n represents the number of heads in n tosses. Put

$$I := \{(m,n) \in \mathbb{Z}^2 : 0 \le m \le n\}.$$

$$\Pi((m,n);(m+1,n+1)) = 1 - \Pi((m,n);(m,n+1)) = \frac{1}{2}$$

Then, for (m,n) and (r,s) in I,

$$\Gamma((m,n);(r,s)) = \begin{cases} \binom{s-n}{r-m} 2^{-(s-n)} & \text{if } 0 \le r-m \le s-n \\ 0 & \text{otherwise} \end{cases}$$

Taking $b = (0,0)$, we find from Stirling's formula that if $s \to \infty$ and $r/s \to t \in [0,1]$ then

$$\kappa((m,n);(r,s)) \to h_t(m,n) := 2^n t^m (1-t)^{n-m}$$

The Martin topology can be regarded as identifying (m, n) in I with $(1+n)^{-1}(m, n) \in \mathbb{R}^2$, with $F \backslash I = [0,1] \times \{1\}$ and

$$h_t = \kappa(\cdot, \xi) \quad (\xi = (t, 1) \in F \backslash I).$$

Thus f is a regular element of S if and only if there exists a probability measure ν on $[0,1]$ such that

$$f(m, n) = \int_0^1 2^n t^m (1-t)^{n-m} \nu(dt) \tag{19}$$

The Weierstrass Approximation Theorem makes it obvious that ν is uniquely determined by f in (19). Hence, h_t is extremal for every $t \in [0,1]$.

Probabilistic interpretation of the Martin compactification

Recall that ζ is the life-time of X.

Theorem 17

Almost surely on $\{\zeta = \infty\}$

$$X_{\zeta-} := \lim_n X_n$$

exists in the topology of F *and* $X_{\zeta-} \in F_e$.

Put

$$X_{\zeta-} := X_{\zeta-1} \text{ on } \{\zeta < \infty\}.$$

(We are not interested in the case when X starts at ∂.) Recall that b is the reference state in terms of which the Martin kernel is defined.

Theorem 18

Let

$$1 = \int_{F_e} \kappa(\cdot, \xi) \, \nu_1(d\xi)$$

be the Martin representation of the (excessive) constant function 1. *Then*

$$\mathbb{P}^b\{X_{\zeta-} \in B\} = \nu_1(B), \quad B \subseteq F_e.$$

To get $f(\cdot,\cdot) = 1$ in the space-time coin-tossing example (19), we have to choose ν to be the unit mass at $\frac{1}{2}$. Thus, Theorems 17 and 18 contain the strong law for tossing a fair coin.

Doob h-transforms

- Assume that $h \in S$ is strictly positive on I.
- The Doob h-transform Π_h of Π is defined as follows:

$$\Pi_h(i,j) := h(i)^{-1}\Pi(i,j)h(j)$$

Then Π_h is sub-stochastic. We have, with obvious notation,

$$\kappa_h(i,j) = \frac{h(b)}{h(i)}\kappa(i,j)$$

and $f \in S_h$ if and only if $hf \in S$.

- Thus, F and F_e are unaffected if we change from Π to Π_h.
- Hence if $X^{(h)}$ is a chain with one-step transition matrix Π_h then

$$X^{(h)}(\zeta^{(h)}-) \text{ exists in } F_e \text{ almost surely.}$$

Here $\zeta^{(h)}$ denotes $\inf\{n : X^{(h)}(n) = \partial\}$.

Theorem 19

A strictly positive function h in S is extremal in S if and only if for some single point ξ of F, we have

$$X^{(h)}(\zeta^{(h)}-) = \xi, \text{ almost surely.}$$

Then, $\xi \in F_e$ and $h = \kappa(\cdot, \xi)$.

We say in this case that $X^{(h)}$ is X conditioned to converge to ξ.

Example: Space-time coin-tossing

Return to space-time coin-tossing and take $h = h_t$ for some fixed $t \in [0, 1]$. Then $h = \kappa(\cdot, \xi)$, where $\xi = (t, 1) \in F \backslash I$, and

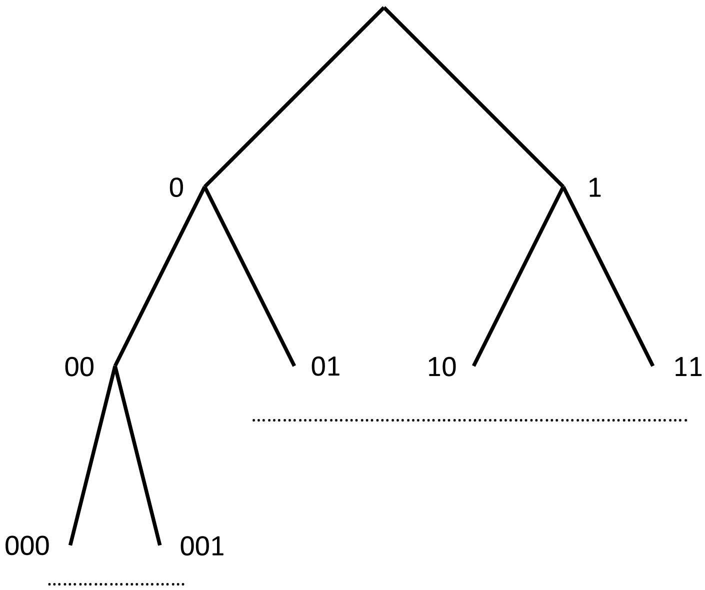
$$\Pi_h((m, n); (m + 1, n + 1)) = 1 - \Pi_h((m, n); (m, n + 1)) = t.$$

Thus, the h-transform corresponds to the case of space-time coin-tossing for a coin with probability t of heads. By the strong law of large numbers, $X^{(h)} \to \xi$ so that h_t is extremal.

- Consider the Pólya sequence $\{X_n\}_{n=1}^{\infty}$ with values in a finite set E for some parameter μ. Put $Y_n := \sum_{k=1}^{n} \delta_{X_k}$ and $Y_0 := 0$.
- Show that the points added in Martin compactification may be identified with space of probability measures on E and the trace of the Martin topology on these points is homeomorphic to the usual topology of weak convergence of probability measures.
- Show that that with this identification, the h-transformed process associated with a probability measure ν is the process constructed in the same manner as Y from an i.i.d. sequence with common distribution ν.

Complete rooted binary tree

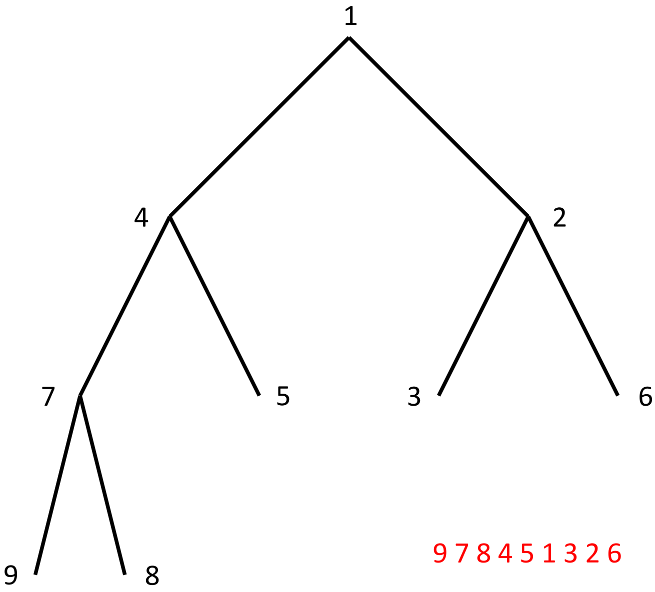
- Denote by $\{0,1\}^\star := \bigsqcup_{k=0}^\infty \{0,1\}^k$ the set of finite tuples or words drawn from the alphabet $\{0,1\}$ (with the empty word $\emptyset$ allowed).
- Write an ℓ-tuple $(v_1, \ldots, v_\ell) \in \{0,1\}^\star$ more simply as $v_1 \ldots v_\ell$.
- Define a directed graph with vertex set $\{0,1\}^\star$ by declaring that if $u = u_1 \ldots u_k$ and $v = v_1 \ldots v_\ell$ are two words, then (u, v) is a directed edge (that is, $u \to v$) if and only if $\ell = k + 1$ and $u_i = v_i$ for $i = 1, \ldots, k$.
- This directed graph is the complete rooted binary tree (rooted at $\emptyset$).


0
1
00
01
10
11
000
001

- A finite rooted binary tree is a non-empty subset $\mathbf{t}$ of $\{0,1\}^\star$ with the property that if $v \in \mathbf{t}$ and $u \in \{0,1\}^\star$ is such that $u \to v$, then $u \in \mathbf{t}$.
- The vertex $\emptyset$ (that is, the empty word) belongs to any such tree $\mathbf{t}$ and is the root of $\mathbf{t}$.

- Suppose that $r(1), \ldots, r(n)$ is a listing of $[n] := \{1, 2, \ldots, n\}$ in some order.
- Label n of the vertices of the complete binary tree as follows:
  - Label $\emptyset$ with 1.
  - Successively for $2 \le k \le n$, attempt to label a vertex with k, starting with the root $\emptyset$. If the vertex u we are currently trying to label with k is already labeled (by some $1 \le j \le k$), next try to label the vertex
    - $u0$ if k is before j in the list,
    - $u1$ if k is after j in the list.

    Stop when we reach an unlabeled vertex.
- The set of labeled vertices form a finite rooted binary tree, and we can recover $r(1), \ldots, r(n)$ from the tree and its labels.


1
4
2
7
5
3
6
9
8
9 7 8 4 5 1 3 2 6

- Let $\{U_k\}_{k=1}^{\infty}$ be i.i.d. uniform r.v. on $[0,1]$.
- Define a random permutation Π_n of $[n]$ by requiring that $\Pi_n(i) < \Pi_n(j)$ if and only if $U_i < U_j$.
- Each permutation Π_n is uniformly distributed.
- Identify Π_n with the ordered list

$$R_n(1), \ldots, R_n(n) = \Pi_n^{-1}(1), \ldots, \Pi_n^{-1}(n).$$

  That is,

$$U_{R_n(1)} < U_{R_n(2)} < \ldots < U_{R_n(n)}.$$

- The corresponding ordered list for Π_{n+1} is obtained by inserting $n+1$ into one of the $n-1$ "slots" between the successive elements of the list or into one of the two "slots" at the beginning and end of the list, with all $n+1$ possibilities being equally likely.

- Apply the procedure for building labeled rooted binary trees to the successive permutations $\Pi_1, \Pi_2, \ldots$ to produce a sequence of labeled trees $\{L_n\}_{n=1}^\infty$, where L_n has n vertices labeled by $[n]$.
- Exercise: Show that the sequence $\{L_n\}_{n=1}^\infty$ is a Markov chain that evolves as follows. Given L_n, there are $n+1$ words of the form $v = v_1 \ldots v_\ell$ such that v is not a vertex of the tree L_n but the word $v_1 \ldots v_{\ell-1}$ is. Pick such a word uniformly at random and adjoin it (with the label $n+1$ attached) to produce the labeled tree L_{n+1}.

- If we remove the labels from each tree L_n, then the resulting sequence $\{T_n\}_{n=1}^{\infty}$ of finite rooted binary trees is also a Markov chain.
- Write G_n (respectively, D_n) for the number of vertices in T_n of the form $0v_2 \ldots v_\ell$ (resp. $1v_2 \ldots v_\ell$). That is, G_n and D_n are, respectively, the sizes of the left and right subtrees below the root $\emptyset$.
- Note that $G_n + 1$ and $D_n + 1$ are, respectively, the number of "slots" to the left and to the right of 1 in the collection of $n+1$ slots between successive elements or at either end of the ordered list $R_n(1), \ldots, R_n(n)$.
- Exercise: Show that $\{(G_n, D_n)\}_{n=1}^{\infty}$ is a Markov chain and conditional on the past up to time n, if $(G_n, D_n) = (g, d)$, then (G_{n+1}, D_{n+1}) takes the values $(g+1, d)$ and $(g, d+1)$ with respective conditional probabilities $\frac{g+1}{g+d+2}$ and $\frac{d+1}{g+d+2}$.
- That is, there is a Pólya sequence $\{X_n\}_{n=1}^{\infty}$ with state-space $E := \{\text{left}, \text{right}\}$ and parameter $\mu(\{\text{left}\}) = \mu(\{\text{right}\}) = 1$ (i.e. a classical Pólya urn) such that

$$G_n := \#\{1 \le k \le n : X_k = \text{left}\}$$

$$D_n := \#\{1 \le k \le n : X_k = \text{right}\}.$$

- For a fixed word $u_1 \dots u_k \in \{0,1\}^\star$, the pair consisting of
  - the number of vertices of the form $u_1 \dots u_k 0 v_2 \dots v_\ell$
  - the number of vertices of the form $u_1 \dots u_k 1 v_2 \dots v_\ell$

  evolves just like $\{(G_n, D_n)\}_{n=1}^\infty$ provided that we only observe the pair at times when it changes state.
- Thus, the process $\{T_n\}_{n=1}^\infty$ may be constructed from an infinite collection of independent, identically distributed Pólya sequences, with one sequence for each vertex of the complete binary tree $\{0,1\}^\star$, by equipping the sequence for each vertex with a clock that depends on the evolution of the sequences associated with vertices that are on the path from the root to the given vertex.

Exercise

- Put $X_n := T_{n+1}$, so that $X_0 = \{\emptyset\}$. As usual, we can think of starting this Markov chain in other states and write $\mathbb{P}^{\mathrm{t}}$ for the corresponding distribution when the initial state is the finite rooted binary tree t.
- Given a tree t and a vertex $u \in \mathrm{t}$, write $\mathrm{t}(u)$ for the vertices that are below u (that is, the vertices v such that the path from the root $\emptyset$ to v passes through u – this includes u itself).
- Show that

$$\mathbb{P}^{\{\emptyset\}}\{X \text{ hits } \mathbf{t}\} = \prod_{u \in \mathbf{t}} (\#\mathbf{t}(u))^{-1}.$$

- Show more generally that for $\mathbf{s} \subseteq \mathbf{t}$

$$\mathbb{P}^{\mathbf{s}}\{X \text{ hits } \mathbf{t}\} = \binom{\#\mathbf{t}}{\#\mathbf{s}}^{-1} \prod_{v \in \mathbf{t} \setminus \mathbf{s}} (\#\mathbf{t}(v))^{-1},$$

- Conclude that the Martin kernel with base point $\{\emptyset\}$ is

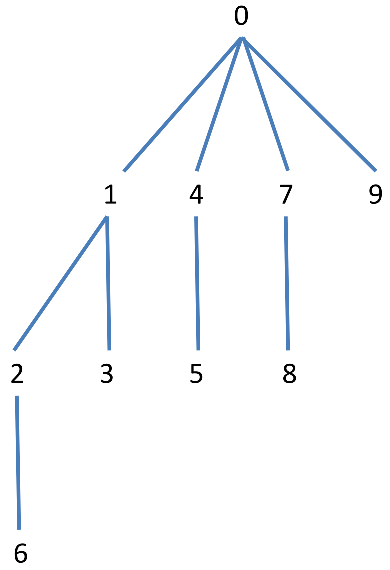
$$K(\mathbf{s}, \mathbf{t}) = \binom{\#\mathbf{t}}{\#\mathbf{s}}^{-1} \prod_{u \in \mathbf{s}} \#\mathbf{t}(u)$$

for $\mathbf{s} \subseteq \mathbf{t}$ (and 0 otherwise).

Another coding of permutations by labeled trees

- A tree is now just a finite, connected, acyclic graph, and the root of such a tree is just a distinguished vertex.
- As before, associate a a permutation π of $[n]$ with the ordered list s$r(1), \ldots, r(n) = \pi^{-1}(1), \ldots \pi^{-1}(n)$.
- The bijection builds a tree with $n+1$ vertices labeled by $[n] \cup \{0\}$.
- The root is labeled with 0.
- If $\pi(i) = \ell$, so that $i = r_\ell$, then the vertex labeled i is the child of the vertex labeled $\max\{r_k : 1 \leq k < \ell, \, r_k < r_\ell\}$, where the maximum of the empty set is 0.
- It is clear that r (equivalently, π) can be reconstructed from the tree and its vertex labels.

9 7 8 4 5 1 3 2 6



- As before, construct random permutations Π_n and random orderings R_n given a sequence $\{U_n\}_{n=1}^{\infty}$ of i.i.d. uniform r.v.
- Obtain random labeled rooted trees L_n and the corresponding random unlabeled rooted trees T_n.
- Both $\{L_n\}_{n=1}^{\infty}$ and $\{T_n\}_{n=1}^{\infty}$ are Markov chains.
- Exercise: Show that given T_n we pick one of its $n+1$ vertices uniformly at random and connect a new vertex to it to form T_{n+1}.
- The chain $\{T_n\}_{n\in\mathbb{N}}$ is the simplest random recursive tree process.

Chinese restaurant connection

- Suppose that the root has k offspring in the tree T_n.
- Write $n_1, \ldots, n_k$ denote the number of vertices in the subtrees rooted at each of these offspring, so that $n_1 + \cdots n_k = n$.
- In constructing T_{n+1} from T_n, either a new vertex is attached to the j^{th} subtree with probability $n_j/(n+1)$ or it is attached to the root and begins a new subtree with probability $1/(n+1)$.
- The manner in which the number and sizes of subtrees rooted at offspring of the root evolve is given by the number and sizes of tables in the simplest Chinese restaurant process: the n^{th} customer to enter the restaurant finds k tables in use with respective numbers of occupants $n_1, \ldots, n_k$ and the the customer either sits at the j^{th} table with probability $n_j/(n+1)$ or starts a new table with probability $1/(n+1)$.
- The random recursive tree process is an infinite hierarchical system of such Chinese restaurant processes.

- Let $\mathbf{I}$ be a countable directed acyclic graph.
- Write $u \to v$ if (u, v) is a directed edge in $\mathbf{I}$.
- Suppose there is a unique vertex $\hat{0}$ such that for any other vertex u there is at least one finite directed path $\hat{0} = v_0 \to v_1 \to \ldots \to v_n = u$ from $\hat{0}$ to u.
- Define a partial order on $\mathbf{I}$ by declaring that $u \leq v$ if $u = v$ or there is a finite directed path $u = w_0 \to w_1 \to \ldots \to w_n = v$.
- Suppose that the number of directed paths between any two vertices is finite.

Parents and children

- For each vertex $u \in \mathbf{I}$, set

$$\alpha(u) := \{v \in \mathbf{I} : v \to u\} = \{\text{immediate predecessors of } u\}$$

and

$$\beta(u) := \{v \in \mathbf{I} : u \to v\} = \{\text{immediate successors of } u\}.$$

- Suppose that $\beta(u)$ is non-empty for all $u \in \mathbf{I}$.

Admissible onward direction tallies

- Suppose for each $u \in \mathbf{I}$ that we have a countable set $\mathbf{S}^u \subseteq (\mathbb{N}_0)^{\beta(u)}$.
- A point $s = (s^v)_{v \in \beta(u)} \in \mathbf{S}^u$ with $\sum_{v \in \beta(u)} s^v = n$ records how many of the first n particles to trickle down through u have been addressed onwards to the various children of u.
- We assume:
  - For each $n \in \mathbb{N}_0$, there exists at least one $s = (s^v)_{v \in \beta(u)} \in \mathbf{S}^u$ such that $\sum_{v \in \beta(u)} s^v = n$. In particular, $(0, 0, \ldots) \in \mathbf{S}^u$.
  - If $t = (t^v)_{v \in \beta(u)} \in \mathbf{S}^u \setminus \{(0, 0, \ldots)\}$, then there exists at least one $s = (s^v)_{v \in \beta(u)} \in \mathbf{S}^u$ such that $t = s + e_v$ for some $v \in \beta(u)$, where e_v is the vector with 1 in the v^{th} coordinate and 0 elsewhere.

- Let $\Sigma^u \subseteq (\mathbf{S}^u)^{\mathbb{N}_0}$ be a set of sequences $(\sigma_n)_{n \in \mathbb{N}_0}$ with the properties:
  - $\sigma_0 = (0, 0, \ldots)$,
  - for each $n \geq 1$, $\sigma_n = \sigma_{n-1} + e_{v_n}$ for some $v_n \in \beta(u)$,
  - each $s = (s^v)_{v \in \beta(u)} \in \mathbf{S}^u$ is of the form σ_n for some $\sigma \in \Sigma^u$, where $n = \sum_{v \in \beta(u)} s^v$.
- The interpretation of $(\sigma_n)_{n \in \mathbb{N}_0}$ is that the n^{th} particle to trickle down to u and find it already occupied is routed onwards to the child v_n.

- Put $\Sigma := \prod_{u \in \mathbf{I}} \Sigma^u$.
- Write a generic element of Σ as

$$((\sigma_n^u)_{n \in \mathbb{N}_0})_{u \in \mathbf{I}}$$

or

$$((\sigma^u(n))_{n \in \mathbb{N}_0})_{u \in \mathbf{I}}.$$

Recall that $\sigma_n^u = \sigma^u(n)$ is an element of $(\mathbb{N}_0)^{\beta(u)}$, and so it has coordinates

$$(\sigma_n^u)^w = (\sigma^u(n))^w, \quad w \in \beta(u).$$

- Fix $\sigma \in \Sigma$.
- Each vertex u of $\mathbf{I}$ has an associated clock $(a_n^u(\sigma))_{n \in \mathbb{N}_0}$ — $a_n^u(\sigma)$ counts the number of particles that have passed through u by time n and been routed onwards to some vertex in $\beta(u)$.
- The family $a_n : \Sigma \to (\mathbb{N}_0)^{\mathbf{I}}$ is defined inductively as follows:
  (a) $a_0(\sigma) = (0, 0, \ldots)$ for all $\sigma \in \Sigma$,
  (b) $a_n^{\hat{0}}(\sigma) = n$ for all $\sigma \in \Sigma$ and $n \in \mathbb{N}_0$,
  (c) $a_n^u(\sigma) = (\sum_{v \in \alpha(u)} (\sigma^v(a_n^v(\sigma)))^u - 1)_+$, $u \neq \hat{0}$.
- The equation in (c) says that the number of particles that have been routed onwards from the vertex u by time n is equal to the number of particles that have passed through vertices v with $v \to u$ and have been routed in the direction of u, excluding the first particle that reached the vertex u and occupied it.

Admissible states

- Let $\mathbf{S}$ denote the subset of $\prod_{u \in \mathbf{I}} \mathbf{S}^u$ consisting of points $x = (x^u)_{u \in \mathbf{I}}$ that can be constructed as $x^u = \sigma^u(a_m^u(\sigma))$ for some $m \in \mathbb{N}_0$ and sequence $\sigma = ((\sigma_n^v)_{n \in \mathbb{N}_0})_{v \in \mathbf{I}} \in \Sigma$.
- Exercise: Check that $(x^u)_{u \in \mathbf{I}} \in \prod_{u \in \mathbf{I}} \mathbf{S}^u$ belongs to $\mathbf{S}$ if and only if

$$\left(\sum_{v \in \alpha(u)} (x^v)^u - 1 \right)_+ = \sum_{w \in \beta(u)} (x^u)^w. \tag{20}$$

Partial order

- Given $x, y \in \mathbf{S}$, say that $x \preceq y$ if for some $m, n \in \mathbb{N}_0$ with $m \leq n$ and some $\sigma = ((\sigma_k^u)_{k \in \mathbb{N}_0})_{u \in \mathbf{I}} \in \Sigma$ we have $x^u = \sigma^u(a_m^u(\sigma))$ and $y^u = \sigma^u(a_n^u(\sigma))$ for all $u \in \mathbf{I}$.
- Exercise: Show that $x \preceq y$ if and only if $(x^u)^v \leq (y^u)^v$ for all $u \in \mathbf{I}$ and $v \in \beta(u)$.

Example

- Suppose that $\mathbf{I}$ is a tree.
- Take $\mathbf{S}^u = (\mathbb{N}_0)^{\beta(u)}$ for all $u \in \mathbf{I}$.
- There is a bijection between $\mathbf{S}$ and finite subtrees of $\mathbf{I}$ that contain the root $\hat{0}$.
  - An element $x \in \mathbf{S}$ is associated with the subtree $\mathbf{t}$ given by

$$\mathbf{t} = \{\hat{0}\} \cup \{v \in \mathbf{I} \setminus \{\hat{0}\} : (x^u)^v > 0 \text{ for some } u \in \alpha(v)\}.$$

  - Conversely, if $\mathbf{t}$ is a finite subtree of $\mathbf{I}$ that contains $\hat{0}$, then the corresponding element of $\mathbf{S}$ is

$$x = \left((\#\{w \in \mathbf{t} : v \leq w\})_{v \in \beta(u)} \right)_{u \in \mathbf{I}}.$$

- The partial order $\preceq$ on $\mathbf{S}$ is equivalent to containment of the associated subtrees.
- The sequences $(x_n)_{n \in \mathbb{N}_0}$ in $\mathbf{S}$ constructed by setting $x_n^u = \sigma^u(a_n^u(\sigma))$ for some $\sigma \in \Sigma$ correspond bijectively to sequences of growing subtrees that begin with the trivial tree $\{\hat{0}\}$ and successively add a single vertex that is connected by a directed edge to a vertex present in the current subtree.

- Assume for each $u \in \mathbf{I}$ that there is a transition matrix Q^u with rows and columns indexed by $\mathbf{S}^u$ such that $\sigma = (\sigma_n)_{n \in \mathbb{N}_0} \in \Sigma^u$ if and only if $Q_u(\sigma_n, \sigma_{n+1}) > 0$ for all $n \in \mathbb{N}_0$.
- Write $(Y_n^u)_{n \in \mathbb{N}_0}$ for the corresponding $\mathbf{S}^u$-valued Markov chain with its associated collection of probability measures $\mathbb{Q}^{u,\xi}$, $\xi \in \mathbf{S}^u$. By assumption, Y^u has positive probability under $\mathbb{Q}^{u,(0,0,\ldots)}$ of hitting any given state in $\mathbf{S}^u$.
- Set

$$A_n := \begin{cases} a_n(Y), & \text{if } Y_0 = (0,0,\ldots), \\ 0, & \text{otherwise.} \end{cases}$$

- Define

$$Z_n^u := Y_{A_n^u}^u, \quad u \in \mathbf{I},\ n \in \mathbb{N}_0.$$

  - By construction, $Z := (Z_n)_{n \in \mathbb{N}_0} = ((Z_n^u)_{u \in \mathbf{I}})_{n \in \mathbb{N}_0}$ is a Markov chain on the countable state space $\mathbf{S}$ under the probability measure $\bigotimes_{u \in \mathbf{I}} \mathbb{Q}^{u,0}$.
  - The paths of Z start from the state $(0,0,\ldots)$ and increase strictly in the natural partial order on $\mathbf{S}$.

By standard arguments, we can construct a measurable space $(\Omega, \mathcal{F})$, a family of probability measures $(\mathbb{P}^x)_{x \in \mathbf{S}}$ and an $\mathbf{S}$-valued stochastic process $X = (X_n)_{n \in \mathbb{N}_0}$ such that X under $\mathbb{P}^x$ is a Markov chain with $X_0 = x$ and the same transition mechanism as Z.

Proposition 20

The Martin kernel of the Markov chain X with respect to the reference state $\hat{0}$ is given by

$$\kappa(x, y) = \prod_{u \in \mathbf{I}} \kappa^u(x^u, y^u),$$

where κ^u is the Martin kernel of the Markov chain Y^u with respect to reference state $(0, 0, \ldots) \in \mathbf{S}^u$. The product is zero unless $x \preceq y$ (equivalently, $x^u \leq y^u$ for all $u \in \mathbf{I}$). Only finitely many terms in the product differ from 1, because $x^u = (0, 0, \ldots)$ for all but finitely many values of $u \in \mathbf{I}$.

Exercise: Show this. Use it to re-compute the Martin kernel for the binary search tree process.

- Recall that we can identify $\mathbf{S}$ for the BST with the set of finite subtrees of the complete binary tree $\{0,1\}^\star$ that contain the root $\emptyset$.
- A sequence $(\mathbf{t}_n)_{n\in\mathbb{N}}$ in $\mathbf{S}$ with $\#\mathbf{t}_n \to \infty$ converges in the Martin compactification of $\mathbf{S}$ if and only if $\#\mathbf{t}_n(u)/\#\mathbf{t}_n$ converges for all $u \in \{0,1\}^\star$. Moreover, if the sequence converges, the limit can be identified with the probability measure μ on $\{0,1\}^\infty$ such that $\mu_u = \lim_{n\to\infty} \#\mathbf{t}_n(u)/\#\mathbf{t}_n$ for all $u \in \{0,1\}^\star$.
- The set of points adjoined to $\mathbf{S}$ in the Martin compactification is homeomorphic to the set of probability measures on $\{0,1\}^\infty$ equipped with the weak topology corresponding to the usual product topology on $\{0,1\}^\infty$.
- Given a probability measure μ on $\{0,1\}^\infty$, write μ_u is the mass assigned by μ to the set of infinite paths in the complete binary tree that begin at the root and that pass through the vertex u. The extended Martin kernel is given by

$$\kappa(\mathbf{s},\mu) = (\#\mathbf{s})! \prod_{u\in\mathbf{s}} \mu_u, \quad \mathbf{s} \in \mathbf{S},\ \mu \in \partial\mathbf{S}. \tag{21}$$

- The transition matrix of the BST is

$$P(\mathbf{s},\mathbf{t}) = \begin{cases} \frac{1}{\#\mathbf{s}+1}, & \text{if } \mathbf{s} \subset \mathbf{t} \text{ and } \#(\mathbf{t} \setminus \mathbf{s}) = 1, \\ 0, & \text{otherwise,} \end{cases}$$

- Set $h_\mu := K(\cdot, \mu)$. The corresponding Doob h-transform process has transition matrix

$$P^{(h_\mu)}(\mathbf{s},\mathbf{t}) = \begin{cases} \mu_u, & \text{if } \mathbf{t} = \mathbf{s} \sqcup \{u\}, \\ 0, & \text{otherwise.} \end{cases}$$

Exercise: Verify that the h-transformed process results from the general trickle-down construction where the routing chains are suitable space-time coin-tossing processes.

Trickle-up construction of the h-transform process

- Suppose on some probability space that there is a sequence of independent identically distributed $\{0,1\}^\infty$-valued random variables $(V^n)_{n \in \mathbb{N}}$ with common distribution μ.
- For an initial finite rooted subtree $\mathbf{w}$,define a sequence $(W_n)_{n \in \mathbb{N}_0}$ of random finite subsets of $\{0,1\}^\star$ inductively by setting $W_0 := \mathbf{w}$ and

$$W_{n+1} := W_n \cup \{V_1^{n+1} \ldots V_{H(n+1)}^{n+1}\}, \quad n \geq 0,$$

where

$$H(n+1) := \max\{h \in \mathbb{N} : V_1^{n+1} \ldots V_h^{n+1} \in W_n\}$$

with the convention $\max \emptyset = 0$.

- That is, at each point in time we start a particle at a "leaf" of the complete binary tree $\{0,1\}^\star$ picked according to μ and then let that particle trickle up the tree until it can go no further because its path is blocked by previous particles that have come to rest.
- Exercise: Show that $(W_n)_{n \in \mathbb{N}_0}$ is a Markov chain with transition matrix $P^{(h_\mu)}$.

- Consider the special case of the h-transform construction when the boundary point μ is the "uniform" or "fair coin-tossing" measure on $\{0,1\}^\infty$; that is, μ is the infinite product of copies of the measure on $\{0,1\}$ that assigns mass $\frac{1}{2}$ to each of the subsets $\{0\}$ and $\{1\}$.
- In this case, the transition matrix of the h-transformed process is

$$P^{(h_\mu)}(\mathbf{s},\mathbf{t}) = \begin{cases} 2^{-|u|}, & \text{if } \mathbf{t} = \mathbf{s} \sqcup \{u\}, \\ 0, & \text{otherwise,} \end{cases}$$

where we write $|u|$ for the length of the word u; that is, $|u| = k$ when $u = u_1 \dots u_k$.

- This transition mechanism is that of the digital search tree process.

Exercise

Analyze the Martin compactification of the random recursive tree process.