Recursive Distributional Equation and Its Applications

Antar Bandyopadhyay

International Conference on the Future of Statistical
Theory, Practice and Education
Hyderabad, India

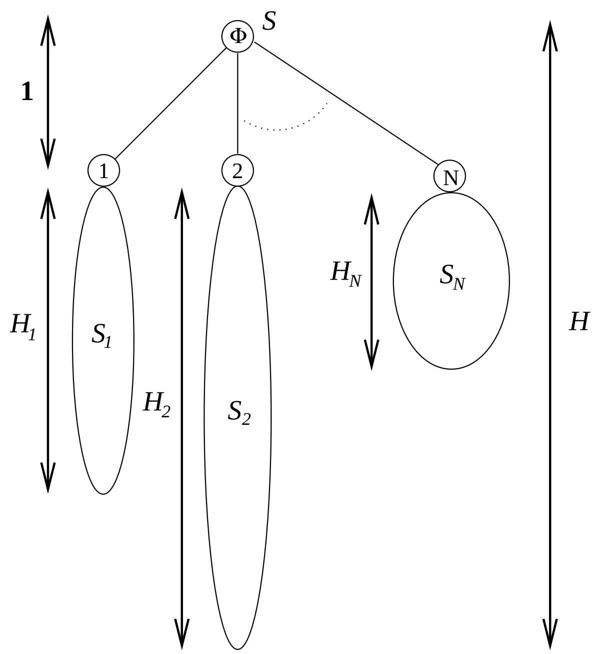
Department of Mathematics
Chalmers University of Technology
Göteborg, Sweden

http://www.math.chalmers.se/~antar

December 31, 2004

Two Examples

Examples 1 : Consider a *(sub)-critical* Galton-Watson branching process with the progeny distribution N, so $\mathbf{E}[N] \leq 1$; we assume $\mathbf{P}(N = 1) < 1$.



Height of the Tree : Let $H := 1+$ height of the G-W tree, then $H < \infty$ a.s. and

$$H \stackrel{d}{=} 1 + \max(H_1, H_2, \ldots, H_N) \quad \text{on } \mathbb{N},$$

where $(H_j)_{j \geq 1}$ are i.i.d. with same law as of H and are independent of N.

Example 2 (Perhaps the best known !) : Consider the following fixed point equation

$$Z \stackrel{d}{=} \frac{Z_1 + Z_2}{\sqrt{2}} \quad \text{on} \quad \mathbb{R},$$

where (Z_1, Z_2) are i.i.d. copies of Z.

- The set of all solutions is given by the Normal $\left(0, \sigma^2\right)$, $\sigma^2 \geq 0$ family.
- This example also extends to give characterizations of stable laws.

We will call such an equation a *recursive distributional equation* (RDE).

Typical features of RDEs

$$\text{Ex. } 1: \quad X \stackrel{d}{=} 1 + \max(X_1, X_2, \ldots, X_N) \quad \text{on } \mathbb{N}$$

$$\text{Ex. } 2: \quad X \stackrel{d}{=} (Z_1 + Z_2)/\sqrt{2} \quad \text{on } \mathbb{R}$$

- **Unknown Quantity :** Distribution of X.
- **Known Quantities :**
  - $N \leq \infty$ which may or may not be random (e.g. $N \equiv 2$ in Ex. 2).
  - Possibly some more randomness whose distribution is known (not present in both examples above).
  - How we combine the known and unknown randomness (e.g. "$1 + \max$" operation in Ex. 1).
- **What is the RDE doing ?** To find a distribution μ such that when we take i.i.d. samples $(X_j)_{j \geq 1}$ from it and only use N many of them (where N is independent of the samples) and do the manipulation then we end up with another sample $X \sim \mu$.

Remark : In the case $N = 1$ a.s. it reduces to the question of finding a stationary distribution of a discrete time Markov chain.

Two main uses of RDEs

- **Direct use :** The RDE is used directly to define a distribution. Examples include,
  - The height (and also the size) of a (sub)-critical Galton-Watson tree (Example 1).
  - The Quicksort distribution (from random algorithm literature).
  - Discounted tree sums / inhomogeneous percolation on trees.
  - ... *and many others*.

- **Indirect use:** The RDE is used to define some auxiliary variables which help in defining/characterizing some other quantity of interest. Among others the following two type of applications are of special interest
  - Characterizing *phase transition* or determining critical points and scaling laws. (will see an example.)
  - 540° *argument* ! (will not give an example.) [Aldous 2000, 2001 and Aldous & B. 2004]

General Setup

- Let $(S, \mathfrak{S})$ be a measurable space, and $\mathcal{P}$ be the collection of all probabilities on $(S, \mathfrak{S})$.

- Let (ξ, N) be a pair of random variables such that N takes values in $\{0, 1, 2, \ldots; \infty\}$.

- Let $(X_j)_{j \geq 1}$ be **i.i.d** S-valued random variables, which are independent of (ξ, N).

- $g(\cdot)$ is a S-valued measurable function with appropriate domain.

Recursive Distributional Equation (RDE)

Definition 1 *The following* fixed-point equation *on* $\mathcal{P}$ *is called a* Recursive Distributional Equation (RDE)

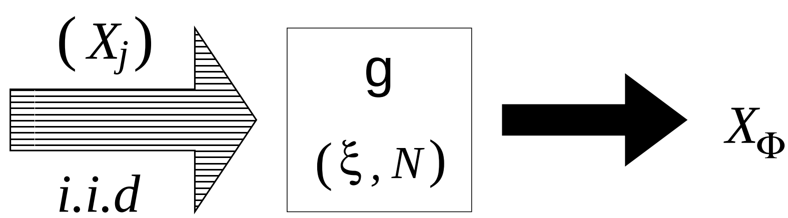
$$X \stackrel{d}{=} g\left(\xi; X_j, 1 \leq j \leq^* N\right), \quad \textit{on} \;\; S$$

where $(X_j)_{j \geq 1}$ *are independent copies of* X *and are independent of* (ξ, N)*.*

Remark : A more conventional (analysis) way of writing the equation would be

$$\mu = T(\mu)$$

where T is the operator associated with the above equation, which depends on the function g and the joint distribution of the pair (ξ, N), and μ is the (unknown) law of X.



Hard-Core Model on a Finite Graph

- Let $G := (V, E)$ be a finite graph.
- We say a subset $I \subseteq V$ is an *independent set* of G, if for any two $u, v \in I$ there is no edge between u and v.
- Let $\mathcal{I}_G$ be the set of all independent sets of G.
- Fix $0 < p < 1$ and let $\lambda = \frac{p}{1-p}$.
- Suppose $(C_v)_{v \in V}$ be i.i.d. Bernoulli$\,(p)$.
- Define $I := \{v \in V \,|\, C_v = 1\}$.
- The measure $\mathbf{P}\left(\cdot \,|\, I \in \mathcal{I}_G\right)$ on $\mathcal{I}_G$ is called the *hard-core model* or *random independent set model* with *activity* λ. We will denote it by $\mathbb{P}$.
- It is easy to see that $\mathbb{P}$ is the probability on $\mathcal{I}_G$ which puts mass proportional to $\lambda^{|I|}$ for $I \in \mathcal{I}_G$.

Sparse Random Graphs

- Two types of *sparse random graphs* :
  - $\mathcal{G}\left(n, \frac{\mu}{n}\right)$: A random graph with n vertices and each edge is present with probability $\frac{\mu}{n}$ independently, where $\mu > 0$. [Erdös & Rényi 1959 - 1968]
  - $\mathcal{G}(n, r+1)$: Pick a graph uniformly at random from the set of all $(r+1)$-regular graphs with n vertices.
- Given a particular realization G_ω of a sparse random graph, we will consider the *hard-core model* with activity $\lambda > 0$ on that finite graph as described before.
- Note there are two stages of randomness and there are two parameters,
  - $\mu > 0$ dealing with the randomness of the graph configuration.
  - $\lambda > 0$ dealing with the randomness of the hard-core model given a configuration.

Motivations

- Interesting from Statistical Physics point of view, well studied for non-random graphs. [Kelley 1985, van den Berg & Steif 1994, Brightwell, Häggström & Winkler 1998, Brightwell & Winkler 1999]

- Has applications in engineering fields, like in *multi-cast networking* problems. [Ramanan et al 2002]

- Conjecture of Aldous [2003] :

  For a sparse random graph if I_n be the *maximal independent set* then

  $$\frac{\mathbf{E}\left[|I_n|\right]}{n} \to c \quad \text{as} \quad n \to \infty,$$

  where c is a constant which depends on the model for the sparse random graph.

Remark : For a hard-core model on a finite graph if we take $\lambda \to \infty$ limit then it concentrate on the maximal independent set(s).

Sparse Random Graphs and GW-Trees

- **Known :** If $\mathcal{G}_n$ be a model for sparse random graph then for "large" enough n "locally it looks like" a (possibly random) rooted tree.
  - ▶ For $\mathcal{G}\left(n, \frac{\mu}{n}\right)$ it is rooted Galton-Watson tree with Poisson (μ) offspring distribution.
  - ▶ For $\mathcal{G}\left(n, r+1\right)$ it is rooted $(r+1)$-regular tree.

- **Conclusion :** So for computing "large" n limit of hard-core model on these kind graphs we need to consider the similar model on respective GW-trees.

- **Problem :** The trees we get may be infinite with positive probability.

- **Solution :** In that case we need to consider Gibbs measure with activity $\lambda > 0$ which has appropriate conditional laws ("DLR condition").

- **Warning :** It is then no longer true that there is only one such measure and we will say that a *phase transition* occurs if there are multiple Gibbs measures for a given activity $\lambda > 0$.

Key Recursion on a Finite Tree

- Suppose $\mathcal{T}$ be a finite rooted tree and we consider the hard-core model on it with activity $\lambda > 0$.

- Suppose $\emptyset$ be the root and it has $n(\emptyset)$ many children which are denoted by $1, 2, \ldots, n(\emptyset)$.

- Let I be a random independent set distributed according to the hard-core model with activity $\lambda > 0$. Then we define $\eta_{\mathcal{T}}^{\emptyset} := \mathbb{P}(\emptyset \in I)$.

- For a child j, let $\mathcal{T}_j$ be the sub-tree rooted at j obtained by removing $\emptyset$. Suppose $\eta_{\mathcal{T}_j}^{j}$ be defined similarly of $\eta_{\mathcal{T}}^{\emptyset}$.

- The following *key recursion* holds

$$\eta_{\mathcal{T}}^{\emptyset} = \frac{\lambda \prod_{j=1}^{n(\emptyset)} \left(1 - \eta_{\mathcal{T}_j}^{j}\right)}{1 + \lambda \prod_{j=1}^{n(\emptyset)} \left(1 - \eta_{\mathcal{T}_j}^{j}\right)}$$

Related RDE

$$\eta \stackrel{d}{=} \frac{\lambda \prod_{j=1}^{N} \left(1 - \eta_j\right)}{1 + \lambda \prod_{j=1}^{N} \left(1 - \eta_j\right)} \quad \text{on} \quad [0,1],$$

where (η_j) are i.i.d. copies of η and are independent of N.

Properties : Let T be the associated operator and $S = T^2$ then

- $T\left(\delta_0\right) = \delta_{\lambda/(1+\lambda)}$.
- $\delta_0 \preccurlyeq T\left(m\right) \preccurlyeq \delta_{\lambda/(1+\lambda)}$, for any probability m on $[0,1]$.
- T is anti-monotone $\Rightarrow$ S is monotone.
- So there exist $m_* \preccurlyeq m^*$ two fixed points of S such that $S^n\left(\delta_0\right) \uparrow m_*$ and $S^n\left(\delta_{\lambda/(1+\lambda)}\right) \downarrow m^*$.
- $T\left(m_*\right) = m^*$.
- S has unique fixed point if and only if $m_* = m^*$.

Uniqueness Domain

Definition 2 *We will say that we are in* uniqueness domain *if* $m_* = m^*$.

Results

- **Theorem 1** *For a GW-Tree with progeny distribution* N *and for activity* $\lambda > 0$ *we are in uniqueness domain if and only if, there is a unique Gibbs measure with activity* λ *a.s. with respect to the randomness in the configuration of the tree.*

  Note : The *phase transition* is characterize by the uniqueness of solution of a RDE.

- **Theorem 2** *For* $\mathcal{G}\left(n, \frac{\mu}{n}\right)$ *suppose we are in the uniqueness domain for* $\lambda > 0$ *and with* $N \sim$ *Poisson* (μ) *and let* I_n *be a random independent set with hard-core distribution with activity* λ*, then*

$$\frac{\mathbf{E}\left[|I_n|\right]}{n} \to \mathbf{E}\left[\eta\right]$$

  where $\eta \sim m_* = m^*$.

- **Theorem 3** *A similar statement for* $\mathcal{G}\left(n, r+1\right)$.

When Uniqueness Domain Holds ?

- **Small μ :** If $\mu \leq 1$ then the graphical structure is in the (sub)-critical domain and hence it will be finite and so uniqueness domain holds for any $\lambda > 0$. This is not the interesting case !

- **Small λ :** If $\lambda \times \mu < 1$ then T is a contraction and hence uniqueness domain holds. Thus for any $\mu > 0$ for activity $\lambda < \frac{1}{\mu}$ we are in the uniqueness domain.

Remarks :

- I believe (do not have complete proofs yet) that uniqueness domain will not hold for large μ or large λ (and $\mu > 1$).

- So it seems that we may not be able to resolve Aldous' conjecture by this method. But perhaps we can ... that is yet another story !

- At least we do get a nice example of *phase transition* phenomenon which is characterize by uniqueness of solution of a RDE.