Modern Discrete Probability

IV - Coupling

Review

Sébastien Roch
UW–Madison
Mathematics

November 29, 2014

Basic definitions

Definition (Coupling)

Let μ and ν be probability measures on the same measurable space $(S, \mathcal{S})$. A *coupling* of μ and ν is a probability measure γ on the product space $(S \times S, \mathcal{S} \times \mathcal{S})$ such that the *marginals* of γ coincide with μ and ν, i.e.,

$$\gamma(A \times S) = \mu(A) \quad \text{and} \quad \gamma(S \times A) = \nu(A), \qquad \forall A \in \mathcal{S}.$$

Similarly, for two random variables X and Y taking values in $(S, \mathcal{S})$, a *coupling* of X and Y is a joint variable (X', Y') taking values in $(S \times S, \mathcal{S} \times \mathcal{S})$ whose law is a coupling of the laws of X and Y. Note that X and Y need not be defined on the same probability space—but X' and Y' do need to.

Examples I

Example (Bernoulli variables)

Let X and Y be Bernoulli random variables with parameters $0 \leq q < r \leq 1$ respectively. That is, $\mathbb{P}[X = 0] = 1 - q$ and $\mathbb{P}[X = 1] = q$, and similarly for Y. Here $S = \{0, 1\}$ and $\mathcal{S} = 2^S$.

- *(Independent coupling)* One coupling of X and Y is (X', Y') where $X' \stackrel{\mathrm{d}}{=} X$ and $Y' \stackrel{\mathrm{d}}{=} Y$ are *independent*. Its law is

$$\left(\mathbb{P}[(X', Y') = (i, j)]\right)_{i,j \in \{0,1\}} = \begin{pmatrix} (1-q)(1-r) & (1-q)r \\ q(1-r) & qr \end{pmatrix}.$$

- *(Monotone coupling)* Another possibility is to pick U uniformly at random in $[0, 1]$, and set $X'' = \mathbb{1}_{\{U \leq q\}}$ and $Y'' = \mathbb{1}_{\{U \leq r\}}$. The law of coupling (X'', Y'') is

$$\left(\mathbb{P}[(X'', Y'') = (i, j)]\right)_{i,j \in \{0,1\}} = \begin{pmatrix} 1-r & r-q \\ 0 & q \end{pmatrix}.$$

Examples II

Example (Bond percolation: monotonicity)

Let $G = (V, E)$ be a countable graph. Denote by $\mathbb{P}_p$ the law of bond percolation on G with density p. Let $x \in V$ and assume $0 \leq q < r \leq 1$.

- Let $\{U_e\}_{e \in E}$ be independent uniforms on $[0, 1]$.
- For $p \in [0, 1]$, let W_p be the set of edges e such that $U_e \leq p$.

Thinking of W_p as specifying the open edges in the percolation process on G under $\mathbb{P}_p$, we see that (W_q, W_r) is a coupling of $\mathbb{P}_q$ and $\mathbb{P}_r$ with the property that $\mathbb{P}[W_q \subseteq W_r] = 1$. Let $\mathcal{C}_x^{(q)}$ and $\mathcal{C}_x^{(r)}$ be the open clusters of x under W_q and W_r respectively. Because $\mathcal{C}_x^{(q)} \subseteq \mathcal{C}_x^{(r)}$,

$$\theta(q) := \mathbb{P}_q[|\mathcal{C}_x| = +\infty] = \mathbb{P}[|\mathcal{C}_x^{(q)}| = +\infty] \leq \mathbb{P}[|\mathcal{C}_x^{(r)}| = +\infty] = \theta(r).$$

Examples III

Example (Biased random walks on $\mathbb{Z}$)

For $p \in [0,1]$, let $(S_t^{(p)})$ be nearest-neighbor random walk on $\mathbb{Z}$ started at 0 with probability p of jumping to the right and probability $1-p$ of jumping to the left. Assume $0 \leq q < r \leq 1$.

- Let $(X_i'', Y_i'')_i$ be an infinite sequence of i.i.d. monotone Bernoulli couplings with parameters q and r respectively.
- Define $(Z_i^{(q)}, Z_i^{(r)}) := (2X_i'' - 1, 2Y_i'' - 1)$.
- Let $\hat{S}_t^{(q)} = \sum_{i \leq t} Z_i^{(q)}$ and $\hat{S}_t^{(r)} = \sum_{i \leq t} Z_i^{(r)}$.

Then $(\hat{S}_t^{(q)}, \hat{S}_t^{(r)})$ is a coupling of $(S_t^{(q)}, S_t^{(r)})$ such that $\hat{S}_t^{(q)} \leq \hat{S}_t^{(r)}$ for all n almost surely. So for all y and all t

$$\mathbb{P}[S_t^{(q)} \leq y] = \mathbb{P}[\hat{S}_t^{(q)} \leq y] \geq \mathbb{P}[\hat{S}_t^{(r)} \leq y] = \mathbb{P}[S_t^{(r)} \leq y].$$

Coupling inequality I

Let μ and ν be probability measures on $(S, \mathcal{S})$. Recall the definition of total variation distance:

$$\|\mu - \nu\|_{\mathrm{TV}} := \sup_{A \in \mathcal{S}} |\mu(A) - \nu(A)|.$$

Lemma

Let μ and ν be probability measures on $(S, \mathcal{S})$. For any coupling (X, Y) of μ and ν,

$$\|\mu - \nu\|_{\mathrm{TV}} \leq \mathbb{P}[X \neq Y].$$

Coupling inequality II

Proof:

$$\begin{aligned}\mu(A) - \nu(A) &= \mathbb{P}[X \in A] - \mathbb{P}[Y \in A] \\ &= \mathbb{P}[X \in A,\ X = Y] + \mathbb{P}[X \in A,\ X \neq Y] \\ &\quad - \mathbb{P}[Y \in A,\ X = Y] - \mathbb{P}[Y \in A,\ X \neq Y] \\ &= \mathbb{P}[X \in A,\ X \neq Y] - \mathbb{P}[Y \in A,\ X \neq Y] \\ &\leq \mathbb{P}[X \neq Y],\end{aligned}$$

and, similarly, $\nu(A) - \mu(A) \leq \mathbb{P}[X \neq Y]$. Hence

$$|\mu(A) - \nu(A)| \leq \mathbb{P}[X \neq Y].$$

■

Example: Poisson distributions

Let $X \sim \mathrm{Poi}(\lambda)$ and $Y \sim \mathrm{Poi}(\nu)$ with $\lambda > \nu$. Recall that a sum of independent Poisson is Poisson. This fact leads to a natural coupling: let $\hat{Y} \sim \mathrm{Poi}(\nu)$, $\hat{Z} \sim \mathrm{Poi}(\lambda - \nu)$ independently of Y, and $\hat{X} = \hat{Y} + \hat{Z}$. Then $(\hat{X}, \hat{Y})$ is a coupling and

$$\|\mu_X - \mu_Y\|_{\mathrm{TV}} \leq \mathbb{P}[\hat{X} \neq \hat{Y}] = \mathbb{P}[\hat{Z} > 0] = 1 - e^{-(\lambda-\nu)} \leq \lambda - \nu.$$

Maximal coupling I

In fact, the inequality is tight. For simplicity, we prove this in the finite case only.

Lemma

Assume S is finite and let $\mathcal{S} = 2^S$. Let μ and ν be probability measures on $(S, \mathcal{S})$. Then,

$$\|\mu - \nu\|_{\mathrm{TV}} = \inf\{\mathbb{P}[X \neq Y] : \textit{coupling } (X, Y) \textit{ of } \mu \textit{ and } \nu\}.$$

Let $A = \{x \in S : \mu(x) > \nu(x)\}$, $B = \{x \in S : \mu(x) \leq \nu(x)\}$ and

$$p := \sum_{x \in S} \mu(x) \wedge \nu(x), \quad \alpha := \sum_{x \in A} [\mu(x) - \nu(x)], \quad \beta := \sum_{x \in B} [\nu(x) - \mu(x)].$$

Maximal coupling II

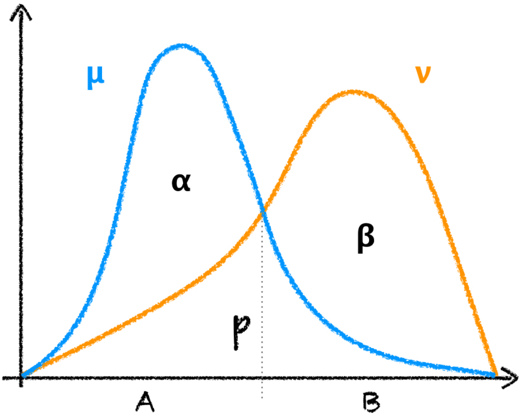

Figure : Proof by picture that: $1 - p = \alpha = \beta = \|\mu - \nu\|_{\mathrm{TV}}$.

Maximal coupling III

Proof: Lemma: $\sum_{x \in S} \mu(x) \wedge \nu(x) = 1 - \|\mu - \nu\|_{\mathrm{TV}}$.
Proof of lemma:

$$\begin{aligned}
2\|\mu - \nu\|_{\mathrm{TV}} &= \sum_{x \in S} |\mu(x) - \nu(x)| \\
&= \sum_{x \in A} [\mu(x) - \nu(x)] + \sum_{x \in B} [\nu(x) - \mu(x)] \\
&= \sum_{x \in A} \mu(x) + \sum_{x \in B} \nu(x) - \sum_{x \in S} \mu(x) \wedge \nu(x) \\
&= 2 - \sum_{x \in B} \mu(x) - \sum_{x \in A} \nu(x) - \sum_{x \in S} \mu(x) \wedge \nu(x) \\
&= 2 - 2 \sum_{x \in S} \mu(x) \wedge \nu(x).
\end{aligned}$$

■

Lemma: $\sum_{x \in A} [\mu(x) - \nu(x)] = \sum_{x \in B} [\nu(x) - \mu(x)] = \|\mu - \nu\|_{\mathrm{TV}} = 1 - p$.
Proof: First equality is immediate. Second equality follows from second line in previous lemma. ■

Maximal coupling IV

The maximal coupling is defined as follows:

- With probability p, pick $X = Y$ from $\gamma_{\min}$ where $\gamma_{\min}(x) := \frac{1}{p}\mu(x) \wedge \nu(x)$, $x \in S$.
- Otherwise, pick X from γ_A where $\gamma_A(x) := \frac{\mu(x)-\nu(x)}{1-p}$, $x \in A$, and, independently, pick Y from $\gamma_B(x) := \frac{\nu(x)-\mu(x)}{1-p}$, $x \in B$. Note that $X \neq Y$ in that case because A and B are disjoint.

The marginal law of X at $x \in S$ is

$$p\gamma_{\min}(x) + (1-p)\gamma_A(x) = \mu(x),$$

and similarly for Y. Finally $\mathbb{P}[X \neq Y] = 1 - p = \|\mu - \nu\|_{\mathrm{TV}}$. ■

Example

Example (Bernoulli variables, continued)

Let X and Y be Bernoulli random variables with parameters $0 \leq q < r \leq 1$ respectively. That is, $\mathbb{P}[X = 0] = 1 - q$ and $\mathbb{P}[X = 1] = q$, and similarly for Y. Here $S = \{0, 1\}$ and $\mathcal{S} = 2^S$. Let μ and ν be the laws of X and Y respectively. To construct the maximal coupling as above, we note that

$$p := \sum_x \mu(x) \wedge \nu(x) = (1 - r) + q, \qquad 1 - p = \alpha = \beta := r - q,$$

$$A := \{0\}, \qquad B := \{1\},$$

$$(\gamma_{\min}(x))_{x=0,1} = \left(\frac{1-r}{(1-r)+q}, \frac{q}{(1-r)+q}\right), \qquad \gamma_A(0) := 1, \qquad \gamma_B(1) := 1.$$

The law of the maximal coupling (X''', Y''') is

$$\left(\mathbb{P}[(X''', Y''') = (i,j)]\right)_{i,j\in\{0,1\}} = \begin{pmatrix} 1-r & r-q \\ 0 & q \end{pmatrix},$$

which coincides with the monotone coupling.

Poisson approximation I

Let $X_1, \ldots, X_n$ be independent Bernoulli random variables with parameters $p_1, \ldots, p_n$ respectively. We are interested in the case where the p_is are "small." Let $S_n := \sum_{i \leq n} X_i$.

We approximate S_n with a Poisson random variable Z_n as follows: let $W_1, \ldots, W_n$ be independent Poisson random variables with means $\lambda_1, \ldots, \lambda_n$ respectively and define $Z_n := \sum_{i \leq n} W_i$. We choose $\lambda_i = -\log(1 - p_i)$ so as to ensure

$$(1 - p_i) = \mathbb{P}[X_i = 0] = \mathbb{P}[W_i = 0] = e^{-\lambda_i}.$$

Note that $Z_n \sim \mathrm{Poi}(\lambda)$ where $\lambda = \sum_{i \leq n} \lambda_i$.

Poisson approximation II

Theorem

$$\|\mu_{S_n} - \mathrm{Poi}(\lambda)\|_{\mathrm{TV}} \leq \frac{1}{2}\sum_{i\leq n}\lambda_i^2.$$

Proof: We couple the pairs (X_i, W_i) independently for $i \leq n$. Let

$$W_i' \sim \mathrm{Poi}(\lambda_i) \quad \text{and} \quad X_i' = W_i' \wedge 1.$$

Because $\lambda_i = -\log(1-p_i)$, (X_i', W_i') is a coupling of (X_i, W_i). Let $S_n' := \sum_{i\leq n} X_i'$ and $Z_n' := \sum_{i\leq n} W_i'$. Then (S_n', Z_n') is a coupling of (S_n, Z_n). By the coupling inequality

$$\|\mu_{S_n} - \mu_{Z_n}\|_{\mathrm{TV}} \leq \mathbb{P}[S_n' \neq Z_n'] \leq \sum_{i\leq n}\mathbb{P}[X_i' \neq W_i'] = \sum_{i\leq n}\mathbb{P}[W_i' \geq 2]$$

$$= \sum_{i\leq n}\sum_{j\geq 2} e^{-\lambda_i}\frac{\lambda_i^j}{j!} \leq \sum_{i\leq n}\frac{\lambda_i^2}{2}\sum_{\ell\geq 0} e^{-\lambda_i}\frac{\lambda_i^\ell}{\ell!} = \sum_{i\leq n}\frac{\lambda_i^2}{2}.$$

■

Maps reduce total variation distance

Theorem

Let X and Y be random variables taking values in $(S, \mathcal{S})$, let h be a measurable map from $(S, \mathcal{S})$ to $(S', \mathcal{S}')$, and let $X' := h(X)$ and $Y' := h(Y)$. It holds that

$$\|\mu_{X'} - \mu_{Y'}\|_{\mathrm{TV}} \leq \|\mu_X - \mu_Y\|_{\mathrm{TV}}.$$

Proof:

$$\begin{aligned}
\sup_{A' \in \mathcal{S}'} \left|\mathbb{P}[X' \in A'] - \mathbb{P}[Y' \in A']\right| &= \sup_{A' \in \mathcal{S}'} \left|\mathbb{P}[h(X) \in A'] - \mathbb{P}[h(Y) \in A']\right| \\
&= \sup_{A' \in \mathcal{S}'} \left|\mathbb{P}[X \in h^{-1}(A')] - \mathbb{P}[Y \in h^{-1}(A')]\right| \\
&= \sup_{A \in \mathcal{S}} \left|\mathbb{P}[X \in A] - \mathbb{P}[Y \in A]\right|.
\end{aligned}$$

■

Erdös-Rényi degree sequence I

Let $G_n \sim \mathbb{G}_{n,p_n}$ be an Erdös-Rényi graph with $p_n := \frac{\lambda}{n}$ and $\lambda > 0$. For $i \in [n]$, let $D_i(n)$ be the degree of vertex i and define

$$N_d(n) := \sum_{i=1}^{n} \mathbb{1}_{\{D_i(n)=d\}}.$$

Theorem

$$\frac{1}{n} N_d(n) \to_{\mathrm{p}} f_d := e^{-\lambda} \frac{\lambda^d}{d!}, \qquad \forall d \geq 1.$$

Proof: We proceed in two steps:

1. we use the coupling inequality to show that the expectation of $\frac{1}{n} N_d(n)$ is close to f_d;
2. we use Chebyshev's inequality to show that $\frac{1}{n} N_d(n)$ is close to its expectation.

Erdös-Rényi degree sequence II

Lemma (Convergence of the mean)

$$\frac{1}{n}\mathbb{E}_{n,p_n}\left[N_d(n)\right] \to f_d, \qquad \forall d \geq 1.$$

Proof of lemma: Note that the $D_i(n)$s are identically distributed so $\frac{1}{n}\mathbb{E}_{n,p_n}\left[N_d(n)\right] = \mathbb{P}_{n,p_n}[D_1(n) = d]$. Moreover $D_1(n) \sim \mathrm{Bin}(n-1, p_n)$. Let $S_n \sim \mathrm{Bin}(n, p_n)$ and $Z_n \sim \mathrm{Poi}(\lambda)$. By the Poisson approximation

$$\|\mu_{S_n} - \mu_{Z_n}\|_{\mathrm{TV}} \leq \frac{1}{2}\sum_{i\leq n}(-\log(1-p_n))^2 = \frac{1}{2}\sum_{i\leq n}\left(\frac{\lambda}{n} + O(n^{-2})\right)^2 = \frac{\lambda^2}{2n} + O(n^{-2}).$$

We can couple $D_1(n)$ and S_n as $(\sum_{i\leq n-1} X_i, \sum_{i\leq n} X_i)$ where the X_is are i.i.d. Bernoulli with parameter $\frac{\lambda}{n}$. By the coupling inequality

$$\|\mu_{D_1(n)} - \mu_{S_n}\|_{\mathrm{TV}} \leq \mathbb{P}\left[\sum_{i\leq n-1} X_i \neq \sum_{i\leq n} X_i\right] = \mathbb{P}[X_n = 1] = \frac{\lambda}{n}.$$

Erdös-Rényi degree sequence III

By the triangle inequality for total variation distance,

$$\frac{1}{2}\sum_{d\geq 0}|\mathbb{P}_{n,p_n}[D_1(n)=d]-f_d| \leq \frac{\lambda+\lambda^2/2}{n}+O(n^{-2}).$$

Therefore,

$$\left|\frac{1}{n}\mathbb{E}_{n,p_n}[N_d(n)]-f_d\right| \leq \frac{2\lambda+\lambda^2}{n}+O(n^{-2}) \to 0.$$

∎

Erdös-Rényi degree sequence IV

Lemma (Concentration around the mean)

$$\mathbb{P}_{n,p_n}\left[\left|\frac{1}{n}N_d(n) - \frac{1}{n}\mathbb{E}_{n,p_n}\left[N_d(n)\right]\right| \geq \varepsilon\right] \leq \frac{2\lambda+1}{\varepsilon^2 n}, \qquad \forall d \geq 1, \forall n.$$

Proof of lemma: By Chebyshev's inequality, for all $\varepsilon > 0$

$$\mathbb{P}_{n,p_n}\left[\left|\frac{1}{n}N_d(n) - \frac{1}{n}\mathbb{E}_{n,p_n}\left[N_d(n)\right]\right| \geq \varepsilon\right] \leq \frac{\mathrm{Var}_{n,p_n}[\frac{1}{n}N_d(n)]}{\varepsilon^2}.$$

Note that

$$\mathrm{Var}_{n,p_n}\left[\frac{1}{n}N_d(n)\right] = \frac{1}{n^2}\left\{\mathbb{E}_{n,p_n}\left[\left(\sum_{i\leq n}\mathbb{1}_{\{D_i(n)=d\}}\right)^2\right] - (n\,\mathbb{P}_{n,p_n}[D_1(n)=d])^2\right\}$$

$$= \frac{1}{n^2}\Big\{n(n-1)\mathbb{P}_{n,p_n}[D_1(n)=d, D_2(n)=d]$$

$$+ n\,\mathbb{P}_{n,p_n}[D_1(n)=d] - n^2\mathbb{P}_{n,p_n}[D_1(n)=d]^2\Big\}$$

Erdös-Rényi degree sequence V

$$\mathrm{Var}_{n,p_n}\left[\frac{1}{n}N_d(n)\right] \leq \frac{1}{n} + \left\{\mathbb{P}_{n,p_n}[D_1(n) = d, D_2(n) = d] - \mathbb{P}_{n,p_n}[D_1(n) = d]^2\right\}$$

We bound the second term using a neat coupling argument. Let Y_1 and Y_2 be independent $\mathrm{Bin}(n-2, p_n)$ and let X_1 and X_2 be independent $\mathrm{Ber}(p_n)$. Then the term in curly bracket above is equal to

$$\begin{aligned}
&\mathbb{P}[(X_1 + Y_1, X_1 + Y_2) = (d,d)] - \mathbb{P}[(X_1 + Y_1, X_2 + Y_2) = (d,d)] \\
&\quad \leq \mathbb{P}[(X_1 + Y_1, X_1 + Y_2) = (d,d),\ (X_1 + Y_1, X_2 + Y_2) \neq (d,d)] \\
&\quad = \mathbb{P}[(X_1 + Y_1, X_1 + Y_2) = (d,d),\ X_2 + Y_2 \neq d] \\
&\quad = \mathbb{P}[X_1 = 0,\ Y_1 = Y_2 = d,\ X_2 = 1] + \mathbb{P}[X_1 = 1,\ Y_1 = Y_2 = d-1,\ X_2 = 0] \\
&\quad \leq \frac{2\lambda}{n}.
\end{aligned}$$

So $\mathrm{Var}_{n,p_n}\left[\frac{1}{n}N_d(n)\right] \leq \frac{2\lambda+1}{n}$. ■ ■

Coupling and bounded harmonic functions I

Lemma

*Let (X_t) be a Markov chain on a (finite or) countable state space V with transition matrix P and let $\mathbb{P}_x$ be the law of (X_t) started at x. Recall that a function $h : V \to \mathbb{R}$ is P-*harmonic *on V (or harmonic for short) if*

$$h(x) = \sum_{y \in V} P(x, y) h(y), \qquad \forall x \in V.$$

If, for all $y, z \in V$, there is a coupling $((Y_t), (Z_t))$ of $\mathbb{P}_y$ and $\mathbb{P}_z$ such that

$$\lim_t \mathbb{P}[Y_t \neq Z_t] = 0,$$

then all bounded harmonic functions on V are constant.

Coupling and bounded harmonic functions II

Proof: Let h be bounded and harmonic on V with $\sup_x |h(x)| = M < +\infty$. Let y, z be any points in V. By harmonicity, $(h(Y_t))$ and $(h(Z_t))$ are martingales and, in particular,

$$\mathbb{E}[h(Y_t)] = \mathbb{E}[h(Y_0)] = h(y) \quad \text{and} \quad \mathbb{E}[h(Z_t)] = \mathbb{E}[h(Z_0)] = h(z).$$

So by Jensen's inequality and the boundedness assumption

$$|h(y)-h(z)| = |\mathbb{E}[h(Y_t)] - \mathbb{E}[h(Z_t)]| \leq \mathbb{E}\,|h(Y_t) - h(Z_t)| \leq 2M\,\mathbb{P}[Y_t \neq Z_t] \to 0.$$

So $h(y) = h(z)$. ■

Harmonic functions on $\mathbb{Z}^d$ I

Theorem

All bounded harmonic functions on $\mathbb{Z}^d$ are constant.

Proof: Clearly, h is harmonic with respect to simple random walk if and only if it is harmonic with respect to lazy simple random walk. Let $\mathbb{P}_y$ and $\mathbb{P}_z$ be the laws of lazy simple random walk on $\mathbb{Z}^d$ started at y and z. We construct a coupling $((Y_t), (Z_t)) = ((Y_t^{(i)})_{i\in[d]}, (Z_t^{(i)})_{i\in[d]})$ of $\mathbb{P}_y$ and $\mathbb{P}_z$ as follows: at time t, pick a coordinate $I \in [d]$ uniformly at random, then

- if $Y_t^{(I)} = Z_t^{(I)}$ then do nothing with probability $1/2$ and otherwise pick $W \in \{-1, +1\}$ uniformly at random, set $Y_{t+1}^{(I)} = Z_{t+1}^{(I)} := Z_t^{(I)} + W$ and leave the other coordinates unchanged;
- if instead $Y_t^{(I)} \neq Z_t^{(I)}$, pick $W \in \{-1, +1\}$ uniformly at random, and with probability $1/2$ set $Y_{t+1}^{(I)} := Y_t^{(I)} + W$ and leave Z_t and the other coordinates of Y_t unchanged, or otherwise set $Z_{t+1}^{(I)} := Z_t^{(I)} + W$ and leave Y_t and the other coordinates of Z_t unchanged.

Harmonic functions on $\mathbb{Z}^d$ II

It is straightforward to check that $((Y_t), (Z_t))$ is indeed a coupling of $\mathbb{P}_y$ and $\mathbb{P}_z$. To apply the previous lemma, it remains to bound $\mathbb{P}[Y_t \neq Z_t]$.

The key is to note that, for each coordinate i, the difference $(Y_t^{(i)} - Z_t^{(i)})$ is itself a random walk on $\mathbb{Z}$ started at $y^{(i)} - z^{(i)}$ with holding probability $1 - \frac{1}{d}$—until it hits 0. Simple random walk on $\mathbb{Z}$ is irreducible and recurrent. The holding probability does not affect the type of the walk, as can be seen for instance from the characterization in terms of effective resistance. So $(Y_t^{(i)} - Z_t^{(i)})$ hits 0 in finite time with probability 1. Hence, letting $\tau^{(i)}$ be the first time $Y_t^{(i)} - Z_t^{(i)} = 0$, we have $\mathbb{P}[Y_t^{(i)} \neq Z_t^{(i)}] \leq \mathbb{P}[\tau^{(i)} > t] \to \mathbb{P}[\tau^{(i)} = +\infty] = 0$.

By a union bound,

$$\mathbb{P}[Y_t \neq Z_t] \leq \sum_{i \in [d]} \mathbb{P}[Y_t^{(i)} \neq Z_t^{(i)}] \to 0,$$

as desired. ■

Harmonic functions on $\mathbb{T}_d$ I

Let $\mathbb{T}_d$ be the infinite d-regular tree with root ρ. For $x \in \mathbb{T}_d$, we let T_x be the subtree, rooted at x, of descendants of x.

Theorem

For $d \geq 3$, let (X_t) be simple random walk on $\mathbb{T}_d$ and let P be the corresponding transition matrix. Let a be a neighbor of the root and consider the function

$$h(x) = \mathbb{P}_x[X_t \in T_a \text{ for all but finitely many } t].$$

Then h is a non-constant, bounded P-harmonic function on $\mathbb{T}_d$.

Harmonic functions on $\mathbb{T}_d$ II

Proof: The function h is clearly bounded and by the usual one-step trick

$$h(x) = \sum_{y \sim x} \frac{1}{d} \mathbb{P}_y[X_t \in T_0 \text{ for all but finitely many } t] = \sum_y P(x,y)h(y),$$

so h is P-harmonic.

Let $b \neq a$ be a neighbor of the root. The key of the proof is:

Lemma

$$q := \mathbb{P}_a[\tau_\rho = +\infty] = \mathbb{P}_b[\tau_\rho = +\infty] > 0.$$

Proof of lemma: Let (Z_t) be simple random walk on $\mathbb{T}_d$ started at a until the walk hits 0 and let L_t be the graph distance between Z_t and the root. Then (L_t) is a biased random walk on $\mathbb{Z}$ started at 1 jumping to the right with probability $1 - \frac{1}{d}$ and jumping to the left with probability $\frac{1}{d}$. The probability that (L_t) hits 0 in finite time is < 1 because $1 - \frac{1}{d} > 2$ when $d \geq 3$. ∎

Harmonic functions on $\mathbb{T}_d$ III

Note that

$$h(\rho) \leq \left(1 - \frac{1}{d}\right)(1 - q) < 1.$$

Indeed if on the first step the random walk started at ρ moves away from a, an event of probability $1 - \frac{1}{d}$, then it must come back to ρ in finite time to reach T_a. Similarly, by the strong Markov property,

$$h(a) = q + (1 - q)h(\rho).$$

Since $h(\rho) \neq 1$ and $q > 0$, this shows that $h(a) > h(\rho)$. ■