Ch5 Coupling

def: A coupling of μ and ν is a pair of random variables (X, Y) s.t. the marginal distribution of X is μ and of Y is ν.

Note: recall $$\|\mu - \nu\|_{TV} = \min_{\substack{(X,Y) \text{ a coupling} \\ \text{of } \mu, \nu}} \Pr(X \neq Y)$$ (Prop 4.7)

Ex: Simple random walk on $\{0, 1, \cdots n\}$ (we go ± 1 unless we are at 0 or n)

We suspect that $x \le y \Rightarrow P^t(x,n) \le P^t(y,n)$

How to prove? Coupling!

Define two walks X_t, Y_t, $X_0 = x$, $Y_0 = y$.

Such that they always both go same direction at each step (if possible).

Note:
- X_t and Y_t almost surely meet up and stay together
- (X_t, Y_t) is a coupling of the distributions $P^t(x, \cdot)$ and $P^t(y, \cdot)$
- $x \le y \Rightarrow X_t \le Y_t \ \forall t$

Therefore $P^t(x,n) = \Pr(X_t = n) \le \Pr(Y_t = n) = P^t(y,n)$ ∎

---

def: A coupling of markov chains w/ trans matrix P is a pair (X_t, Y_t) s.t. both (X_t) and (Y_t) are MC's with matrix P.

Notation: if (X_t, Y_t) is a coupling of MC's with P (and $X_0 = x, Y_0 = y$) then we write $\Pr_{x,y}$ for probabilities on the common space.

Thm: let (X_t, Y_t) be a coupling such that

$$\star \quad X_s = Y_s \implies X_{s+1} = Y_{s+1}$$

and $X_0 = x$, $Y_0 = y$. define $\tau_{couple} := \min\{t : X_t = Y_t\}$

then $\| P^t(x,\cdot) - P^t(y,\cdot)\|_{TV} \leq Pr_{x,y}(\tau_{couple} > t)$.

Pf: Note $P^t(x,z) = Pr_{x,y}(X_t = z)$, $P^t(y,z) = Pr_{x,y}(Y_t = z)$

so (X_t, Y_t) is a coupling of $P^t(x,\cdot)$ and $P^t(y,\cdot)$

Now $\| P^t(x,\cdot) - P^t(y,\cdot)\|_{TV} \leq Pr_{x,y}(X_t \neq Y_t) = Pr_{x,y}(\tau_{couple} > t)$ ∎

def: Markovian Coupling, p. 65

Cor: $$d(t) \leq \max_{x,y \in \Omega} Pr_{x,y}(\tau_{couple} > t)$$

Pf: $d(t) \leq \bar{d}(t) = \max_{x,y} \| P^t(x,\cdot) - P^t(y,\cdot)\|_{TV} \leq$ RHS

Ex: Mixing time for lazy random walk on cycle:

Use the coupling (X_t, Y_t) with $X_0 = x$, $Y_0 = y$ where they never move simultaneously (i.e. either X moves or Y moves)

And let D_t be the clockwise distance from X_t to Y_t.

X_t meets Y_t when $D_t = 0$ or n and then run X_t and Y_t together

Note that D_t is a random walk (non lazy) on $\{0, \cdots n\}$

and $d(t) \le \max_{x,y} Pr_{x,y}(\tau_{couple} > t) \le \frac{\max_{x,y} E_{x,y}(\tau_{couple})}{t}$

but $E_{x,y}(\tau_{couple}) = k(n-k)$ where $D_0 = k$

so $d(t) \le \frac{n^2}{4t}$ and $t_{mix} \le n^2$.

<u>Ex</u>: random lazy walk on torus (describe torus, p 65)

<u>Thm</u>: for the lrw on $\mathbb{Z}_n^d$, $t_{mix}(\varepsilon) \le C(d)\, n^2 \log_2(\varepsilon^{-1})$

where $C(d)$ is a constant depending on d.

<u>Pf</u> (by coupling): Lets couple. $\vec{X}_0 = \vec{x}$, $\vec{Y}_0 = \vec{y}$.

To move: randomly pick a coordinate (of the d dimensions)

If $\vec{X}_t$ and $\vec{Y}_t$ agree in this coordinate, move (or stay) together.

If not, flip a coin to see whether $\vec{X}$ moves or $\vec{Y}$ moves

and flip again to see which direction.

define $\tau_i := \min\{ t \ge 0 : X_t^i = Y_t^i \}$

Coordinate i is selected with probability $\frac{1}{d}$, so the expected wait time between picking i again is d.

$\boxed{\text{Ex 5.3: if } X_1, X_2, \ldots \text{ are ind. with mean } \mu \text{ and } \tau \text{ is } \mathbb{Z}^+\text{-valued independent of the } X\text{'s, then } E\left(\sum_{i=1}^{\tau} X_i\right) = \mu E(\tau)}$

$\Rightarrow E_{x,y}(\tau_i) \le \frac{dn^2}{4}$.

Now $E_{x,y}(\tau_{couple}) = E_{x,y}(\max_i \tau_i) \le \frac{d^2 n^2}{4}$

Now $\Pr_{x,y}(\tau_{couple} > t) \le \frac{E_{x,y}(\tau_{couple})}{t} \le \frac{d^2 n^2}{4t} \Rightarrow t_{mix} \le d^2 n^2$

so $t_{mix}(\varepsilon) \le \lceil \log_2 \varepsilon^{-1} \rceil t_{mix}$ (4.36) ■

LRW on hypercube $\mathbb{Z}_2^n$

Let $\vec{X}_0 = \vec{x}$, $\vec{Y}_0 = \vec{y}$ and couple as follows:

Pick a random coordinate and "refresh" both X and Y in that coordinate by replacing both with a 0 or 1, chosen by a coin toss.

$\tau_{couple} \le \min\{t \ge 0 : \text{all coordinates have been picked}\}$

Now recall Section 2.2, coupon collector problem.

$d(n\log n + cn) \le \Pr(\text{there are still any unpicked coordinates at step } n\log n + cn) \le e^{-c}$

this immediately gives $t_{mix}(\varepsilon) \le n\log n + \log\left(\frac{1}{\varepsilon}\right) n$

distance between $P^t(x,\cdot)$ and $P^{t+1}(x,\cdot)$ for lazy P.

Prop: let Q be irreducible trans. matrix and $P = \frac{1}{2}(I+Q)$.

then $\|P^t(x,\cdot) - P^{t+1}(x,\cdot)\|_{TV} \leq \frac{12}{\sqrt{t}}$

Pf: we construct (X_t), (Y_t) both started at x with trans matrix P s.t. $\Pr(X_t \neq Y_{t+1}) \leq \frac{12}{\sqrt{t}}$

and this suffices since (X_t, Y_{t+1}) is a coupling of the distributions $P^t(x,\cdot)$ and $P^{t+1}(x,\cdot)$.

Let (Z_t) be a MC with matrix Q started at x, and (W_t) an i.i.d sequence of 0-1 valued random variables.

Let $N_t := \sum_{s=1}^{t} W_s$ and $Y_t = Z_{N_t}$.

Note that Y_t is an MC with matrix P.

Let $X_t := \begin{cases} Z_{t-(N_{t+1}-W_1)} & \text{if } X_{t-1} \neq Y_t \\ Y_{t+1} & \text{otherwise} \end{cases}$

X_t is also a MC with matrix P.

Let $\tau := \min\{t \geq 0 : X_t = Y_{t+1}\}$.

Case 1: $W_1 = 0$: then $Y_1 = X_0 = x$ so $\tau = 0$

Case 2: $W_1 = 1$ then $\tau \leq \min\{t \geq 0 : N_{t+1} = t - (N_{t+1} - W_1)\}$

and $N_{t+1} = t - (N_{t+1} - W_1) \Leftrightarrow 2N_{t+1} - t = 1$

but $(2N_{t+1} - t)_{t=0}^{\infty}$ is a random walk starting at 2 but th 2.17 $\Rightarrow$ $\Pr(\tau > t) \le 12/\sqrt{t}$

Grand Coupling

<u>def</u>: a collection $\{ X_t^x : x \in \Omega, \ t = 0, 1, \dots \}$ of random variables such that for each $x \in \Omega$, (X_t^x) is a MC with matrix P started at x is called a Grand Coupling.

<u>Note</u>: can construct using random mapping model:

let $f : \Omega \times \Lambda \to \Omega$ and Z a Λ-valued random var such that $P(x,y) = \Pr(f(x,Z) = y)$. (f, Z) exists by Prop 1.5

Now let $Z_1, Z_2, \dots$ be i.i.d copies of Z and

$X_0^x := x$, $X_t^x = f(X_{t-1}^x, Z_t)$ for $t \ge 1$

Recall the metropolis chain on colorings.

$\tilde{\Omega}$ = Space of colorings, improper allowed

<u>Thm</u>: let $|G| = n$, max degree Δ. If $q > 3\Delta$ and

$C_{met}(\Delta, q) = 1 - \frac{3\Delta}{q}$ then for the metropolis chain on G

$$t_{mix}(\varepsilon) \le C_{met}(\Delta,q)^{-1}\, n\left[\log n + \log \frac{1}{\varepsilon}\right]$$

Pf: use a grand coupling. Keep track of the MC for every possible start coloring in $\tilde{\Omega}$.

def: for $x, y \in \tilde{\Omega}$, $\rho(x,y) := \sum_{v \in V} \mathbb{1}_{\{x(v) \ne y(v)\}}$

Note: ρ is a metric.

Suppose $\rho(x,y) = 1$ so x and y disagree only at v_0.
Let $N = x(N(v_0))$, the set of colors of neighbors of v_0

Then $\Pr\left[\rho(X_1^x, X_1^y) = 0\right] = \frac{1}{n} \cdot \frac{q - |N|}{q} \ge \frac{q - \Delta}{nq}$

There is a case where $\rho(X_1^x, X_1^y) = 2$ (See p. 72)

but $\Pr\left[\rho(X_1^x, X_1^y) = 2\right] \le \frac{\Delta}{n} \cdot \frac{2}{q}$

Thus $\mathbb{E}\left[\rho(X_1^x, X_1^y) - 1\right] \le \frac{3\Delta - q}{nq}$

So $\mathbb{E}\left[\rho(X_1^x, X_1^y)\right] \le 1 - \frac{C_{met}(\Delta,q)}{n}$

Now if $\rho(x,y) = r$ then $\exists$ a sequence

$x = x_0, x_1, \cdots x_r = y$ s.t. $\rho(x_i, x_{i+1}) = 1$

so Δ-ineq $\Rightarrow$ $\mathbb{E}\left[\rho(X_1^x, X_1^y)\right] \le \mathbb{E}\left[\sum_{k=1}^{r} \rho\left(X_1^{x_k}, X_1^{x_{k+1}}\right)\right]$

$\le \rho(x,y) \cdot \left(1 - \frac{c_{met}}{n}\right)$

Now $\mathbb{E}\left[\rho(X_t^x, X_t^y) \mid X_{t-1}^x = x_{t-1}, X_{t-1}^y = y_{t-1}\right] = \mathbb{E}\left[\rho\left(X_1^{x_{t-1}}, X_1^{y_{t-1}}\right)\right]$

$\le \rho(x_{t-1}, y_{t-1}) \left(1 - \frac{c_{met}}{n}\right)$

so $\mathbb{E}\left[\rho(X_t^x, X_t^y)\right] \le \mathbb{E}\left[\rho(X_{t-1}^x, X_{t-1}^y)\right] \left(1 - \frac{c_{met}}{n}\right)$

$\le \cdots \le \rho(x,y) \left(1 - \frac{c_{met}}{n}\right)^t$

Now $\Pr\left[X_t^x \ne X_t^y\right] = \Pr\left[\rho(X_t^x, X_t^y) \ge 1\right] \le \rho(x,y)\left(1 - \frac{c_{met}}{n}\right)^t$

$\le n e^{-\frac{t c_{met}}{n}}$

thus $d(t) \le n e^{-\frac{t c_{met}}{n}}$ and the result follows $\blacksquare$