DTMC and Applications

<u>Def:</u> $(X_n)_{n\geq 0}$ is DTMC if it satisfies Markovian property: $P(X_{n+1}=i_{n+1} \mid X_k = i_k,\ 1\leq k\leq n) = P(X_{n+1}=i_{n+1} \mid X_n = i_n)$

<u>Rmk:</u> i) It's equi.: $P(X_{n+1}\in A_{n+1} \mid X_n = i_n,\ X_k\in A_k,\ k\leq n-1)$

$= P(X_{n+1}\in A_{n+1} \mid X_n = i_n)$. $\forall A_k \subset S$.

(But not true for "i_n" replaced by "A_n")

ii) A: past. B: present. C: future.

$$\Rightarrow \begin{cases} P(C\mid BA) = P(C\mid B) \\ P(AC\mid B) = P(A\mid B)\,P(C\mid B) \end{cases}$$

Next. we consider Time-homogeneous DTMCs:

It's specified by: i) State space S. $|S|\leq$ S.

ii) Initial dist. λ_0. $X_0\sim\lambda_0$.

iii) Prob. transition rules $P=(P_{ij})_{S\times S}$.

<u>Thm.</u> $X=(X_n)_{n\in\mathbb{Z}_{\geq 0}}$ $\iff$ $\exists\ (\xi_n)$ i.i.d. $f:\mathbb{R}^d\times S\to S$. measurable. s.t. $X_{n+1} = f(\xi_{n+1}, X_n)$. where X_0 indept of (ξ_n). Besides. $P_{ij} = P(f(\xi_1, i)=j)$.

Construction: $S = \mathbb{Z}_{\geq 0}$. $\mu = (\mu_k)_S$ $(P_{ij})_{S\times S}$ are dist. and trans. prob. Then $\exists\ (X_n)$. determined by them.

Pf: 1°) Construct $U_k \overset{\text{i.i.d}}{\sim}$ Uniform$[0,1]$.

consider $\bigotimes_{N} m|_{[0,1]}$. m is Lebesgue measure.

2°) Set: $h(u) = \inf\{l : \sum_0^l \mu_k > u\}$.

$f(i,u) = \inf\{l : \sum_0^l P_{ik} > u\}$.

$X_0 = h(U_0)$. $X_n = f(\xi_n, X_{n-1})$.

Thm. $X = (X_n)$ satisfies Strong Markov property:

$\forall\ \tau$ stopping time. w.r.t $\mathcal{F}_n^X$. $\forall\ C \in \mathcal{F}_\tau$. then:

$P(X_{\tau+n} = j \mid C,\ X_\tau = i.\ \tau = l) = p^n(i,j)$.

Pf: Note: $C \cap \{\tau = l\} = (X_0 \cdots X_l) \in A$

check: $P(X_{\tau+n} = j.\ X_\tau = i,\ \tau = l.\ C) = p^n(i,j)\, P(\square)$.

Cor. $(X_{\tau+n})$ is Markov Chain under $\{\tau < \infty\}$.

Pf: $P(C.\ X_{\tau+n} = j.\ X_\tau = i.\ \tau < \infty) =$

$\sum_{\mathbb{Z}_{\geq 0}} P(C.\ X_{\tau+n} = j.\ X_\tau = i.\ \tau = k) =$

$\sum P(C,\ X_k = i.\ \tau = k)\, P_{ij}^n =$

$P(C.\ X_\tau = i,\ \tau < \infty)\, P_{ij}^n$.

Take $C = \Omega$. $\Rightarrow P(X_{\tau+n} = j \mid X_\tau = i.\ \tau < \infty) = \hat{P}_{ij}$.

Def: State $i \in S$ is essential. if $\forall j$.

st. $i \to j \Rightarrow j \to i$.

RMK: i.e. It will return to i. (loop)

prop. i is essential. And $i \to j \Rightarrow j$ is essential.

Pf: $i \to j \to k \to i \to j$. $\forall k \in S$. st. $j \to k$.

Def: i) i is essential. Define $C(i) = \{i\} \cup \{i \to j\}$

is essential class w.r.t. i.

ii) If i is essential. $C(i) = \{i\}$. Then we say i is absorbing.

Thm. $S = T \cup (\cup R_j)$. where T is set of states which're not essential. R_j is essential class. $j \geq 1$

① Criteria of Recurrent:

Thm. $\forall j \in S$. $(e_{ij})_{i \in S}$ is nonnegative solution

of $Z_i = \Sigma P_{ik} Z_k + P_{ij}$. $i \in S$.

Pf: By first step analysis

Thm. $e_{ij}^{(n)} = P_i(N_j \geq n) = e_{ij} e_{jj}^{n-1}$.

Pf: Apply Strong Markov property.

i) Min nonnegative solution:

Thm: $(e_{ij})_{i\in S}$ is minimal negative solution for

$$z_i = \sum_{k\neq j} p_{ik} z_k + p_{ij}. \quad i \in S. \quad \text{for } \forall j.$$

Pf: $\forall (u_i)_S$ is solution of the equation:

Induct on n, prove:

$$u_i \geq \sum_{k=1}^{n} P_i(T_j = k). \quad \forall n. \text{ Set } n \to \infty.$$

$$n=1: \ u_i = \sum_{k\neq j} p_{ik} u_k + p_{ij}$$

$$\geq p_{ij} = P_i(T_j = 1).$$

For $n=k$. replace u_i by inequi. of $n=1$. in RHS. iteritation.

Thm. X is irred. Then. it's recurrent $\Leftrightarrow \exists j \in S$. st.

Any solution of $z_i = \sum_{k\in S} p_{ik} z_k$. $\forall i \neq j$.

which is bdd will be const. (i.e. $z_i = z_j$).

Pf: $(\Leftarrow)$ $(z_k = e_{kj})_{k\neq j} \cup (z_j = 1)$. is

a solution. So: $e_{kj} = 1$. $\forall k \neq j$.

$$e_{jj} = \sum p_{ik} e_{kj} + p_{ij} = 1.$$

ii) Harmonic Function:

Def: i) Solution of $\sum_{k \in S} P_{ik} Z_k \leq Z_i$. $i \in S$

is superharmonic function of X.

ii) If (Z_i) satisfies equation above.

Then we call it's harmonic function

of X.

Thm. For X is irred. Then it's recurrent

$\Leftrightarrow$ $\forall$ nonnegative superharmonic func. is const.

(i.e. in sense of $Z_i = Z_j$. $\forall i, j$).

iii) Foster-Lyapunov Criterion:

Def: $A = P - I$ on $\{ V : S \to \mathbb{R}^1$. measurable $\}$.

is an operator defined by:

$AV(i) =: PV(i) - V(i) =: \sum_{j \in S} P_{ij} V(j) - V(i)$

where $P|V|(i) = \sum P_{ij} |V(j)| < \infty$. $\forall i$.

Thm. For X is irred. If $\exists j \in S$. and $V \geq 0$

on S. st. $AV(i) \leq 0$. $\forall i \neq j$. and

$\forall r < \infty$. $\{ i \mid V(i) \leq r \}$ is finite set.

(WLOG. suppose $S \subset \mathbb{Z}^+$).

Then: X is recurrent.

Rmk: It requires $V \to \infty$. We can see V as some kind of energy function.

③ Criteria for Positive Recurrent:

Thm X is irred. recurrent. For $j \in S$. Then $(\eta_k) = (E_k(T_j))_{k \in S}$ is min nonnegative solution of: $\begin{cases} z_j = 0 \\ \sum_S p_{ik} z_k = z_i - 1, \ i \neq j \end{cases}$ satisfies:

$\sum p_{ik} \eta_k = \eta_i - 1. \ \eta_j = 0. \ \forall i \neq j.$

Pf: By First step analysis.

it's easy to see the equation holds.

Minimal is similar as before:

$u_i \geq 1 + \sum p_{ik} u_k \geq 1 + \sum_{k \neq j} p_{ik} = 1 + P_i(X_1 \neq j)$.

(Note that u_k require: $u_k \geq 1$).

replace in RHS:

$u_i \geq 1 + \sum p_{ik}(1 + P_i(X_1 \neq j)) = 1 + P_i(X_1 \neq j) + P(\cdots)$

Repeat the process:

$u_i \geq 1 + \sum_{k=1}^{n} P(X_1 \sim X_k \neq j) = \sum P_i(T_j \geq m)$

$\to E_i(T_j)$. as $n \to \infty$.

Rmk: For positive recurrence: Prove (z_j) will be bdd where (z_j) is its solution.

Thm: (Foster-Lyapunov)

X is irred. recurrent. Then: X is positive recurrent $\iff$ $\exists\, j \in S$. and $V \geq 0$ on S. $b < \infty$. s.t. $A V(i) \leq -1 + b\,\delta_{ij}$. $i \in S$.

Def: X is uniformly recurrent. if $\sup\limits_{S} E_i(T_i) < \infty$.

Thm. (Foster-Lyapunov)

X is irred. recurrent. Then: X is uniformly recurrent $\iff$ $\exists\, j \in S$. $V \geq 0$ on S. bdd. and $b < \infty$. s.t. $A V(i) \leq -1 + b\,\delta_{ij}$. $\forall\, i \in S$.

③ **Law of Large Number:**

Lemma $Y_m =: \sum\limits_{T_i(m-1)+1}^{T_i(m)} f(X_k)$. $i \in S$ recurrent.

Then, $((T_i(m) - T_i(m-1), Y_m))_{m \geq 1}$ i.i.d. under P_i.

Pf: By strong markov property.

Thm. (SLLN of Markov chain)

X is irred. positive recurrent. If: $f: S \to \mathbb{R}^1$. s.t. $\Sigma_S |f(j)|\, \pi_j < \infty$. Then:

$\forall\, i \in S$. $\lim\limits_{n} \frac{1}{n} \sum\limits_{1}^{n} f(X_m) = \Sigma_S\, \pi_j f(j)$. P_i-a.s.

Pf: By Lemma. SLLN. ERT.

④ **Doeblin Theory:**

Next, consider X is irred. aperiodic. Markov Chain.

Def: X satisfies Doeblin condition if $\exists\ (\mu_j)_S$ prob dist. $r \in (0,1)$. $n_0 \geq 1$. St. $P^{n_0}_{ij} \geq r\mu_j$. $\forall i, j$.

Rmk: By C-K equation:

$$\forall n \geq n_0.\ P^n_{ij} \geq \sum_k P^{n-n_0}_{ik} r\mu_j \geq r\mu_j.$$

Thm. If X satisfies Doeblin condition. Then: $\exists M < \infty$. $\ell \in (0,1)$. St. for $\forall n \geq 1$: We have,

$$\ell_n =: \sup_{i,\ell \in S} \sum_{j \in S} | P^n_{ij} - P^n_{\ell j} | \leq M\ell^n.$$

Def: X is said to be uniformly ergodic if $\forall j \in S$.

$$\lim_{n\to\infty} \sup_i | P^n_{ij} - \pi_j | = 0.$$

Thm. i) X is uniformly ergodic

ii) X satisfies Doeblin condition

iii) $\exists\ (\pi_j)_S$ prob. dist. $M < \infty$. $\ell \in (0,1)$. St. $\forall n \geq 1$.

$$\sup_i \sum_{j \in S} | P^n_{ij} - \pi_j | \leq M\ell^n.$$

We have: i). ii). iii) equi.

(1) **Metroplis Method:**

To simulate a dist. π on S st. $|S| < \infty$.

We can use limit Thm to approxi. it:

Consider on connected graph. Denote $N(j)$ is set of neighbour points of j. $d(i)$ is $\# N(i)$

① Run a random Walk on the graph:

$$P_{rw}(i,j) = \begin{cases} 1/d(i) \,. & j \in N(i) \\ 0 \,. & j \notin N(i) \end{cases}$$

$\pi_{rw}(i) = d(i) / \sum_{j \in S} d(j)$, is stationary dist.

② <u>Modification</u>:

Note $\pi_{rw}(i) \propto d(i)$. if we want: $\pi(i) \propto f(i)\, d(i)$

$$\text{Set}: \ p(i,j) = \begin{cases} \min\{1, f(j)/f(i)\}/d(i) \,. & j \in N(i) \\ 1 - \sum_{j \in N(i)} p(i,j) \,. & j = i \\ 0 \,. & \text{otherwise} \end{cases}$$

<u>Pf</u>: $\pi(i)\, p(i,j) \propto \min\{f(i), f(j)\}$. By symmetry:

$\pi(i)\, p(i,j) = \pi(j)\, p(j,i) \Rightarrow \pi P = \pi$. stationary.

<u>Interpret</u>: 1) Choose a neighbour of i w.p. $1/d(i)$

2) If $f(j) \geq f(i)$. Then accept j.

3) If $f(j) < f(i)$. Then reject w.p. $1 - \frac{f(j)}{f(i)}$

<u>Rmk</u>: We don't need to know $(\pi(i))$. exactly.

Avoid Normalization.

(2) <u>Simulated Annealing</u>:

Target: Find $i \in S$. s.t. $c(i) = \min_{S} c(k)$. where $c(i)$ is the cost. function. def on set of nodes S in a graph. $|S| < \infty$.

Def: prob. dist. $\alpha_T = \{\alpha_T(i) \mid i \in S\}$ on S is:

$$\alpha_T(i) = d(i) e^{-c(i)/T} / G(T). \quad G(T) = \sum_{S} d(i) e^{-c(i)/T}$$

Rmk: i) Choose $f(i) = e^{-c(i)/T}$ in Metroplis Metod

$\pi(i) \propto d(i) f(i) \Rightarrow \pi(i) = \alpha_T(i)$.

ii) T stands for temperature.

Denote: $S^* = \{i^* \in S \mid c(i^*) = \min_{S} c(k)\}$. With a dist. α^* on S^* is: $\alpha^*(i) = d(i) / \sum_{S^*} d(i)$. if $i \in S^*$.

$\alpha^*(i) = 0$. if $i \notin S^*$.

Rmk: We only put positive prob. on optimals.

Prop. $\alpha_T \xrightarrow{d} \alpha^*$. as $T \downarrow 0$.

Pf: Check $\alpha_T(i) \to \alpha^*(i)$: divide $e^{-\min c(i)/T}$. Set $T \downarrow 0$

procedure: 1°) Simulate the stationary α_{T_1} by Metroplis Method. $A_{T_1} = (A_{T_1}(i,j))_{S \times S}$ is its prob. transition matrix. def by (1) ⊙.

2°) When dist. is approaching α_{T_1} nearly. Decrease the temperature to $T_2 < T_1$.

Then approxi. π_{T_2} by A_{T_2} prob. matrix.

3') Repeat these steps. Let $T_n \to 0$ $(n \to \infty)$

Next. Suppose we have $X_0 \sim \nu_0$. We will cool the temperature at each step. i.e. $P(X_n = i \mid X_{n-1} = j) = A_{T_{n-1}}(j, i)$

i.e. $X_n \sim \nu_0 A_{T_0} A_{T_1} \cdots A_{T_{n-1}}$ $T_n \downarrow 0$.

<u>Denote:</u> $A^{(n)} = \prod_{0}^{n-1} A_{T_k}$. $\nu^{(n)} = \nu_0 A^{(n)}$.

$r = \min_{i \in S_c} \max_{j \in S} \ell(i, j)$, ℓ is distance. and denote:

$S_c = \{ i \mid c(j) > c(i),$ exist some $j \in N(i) \}$

$L = \max_{i \in S} \max_{j \in N(i)} | c(j) - c(i) |$. max local fluctuation.

<u>Rmk:</u> S_c is set of points which is not local max. Then r is min radius of $i \in S_c$. st. contain all $j \in S$

$\Rightarrow$ Our goal is simulating π^*:

<u>Thm.</u> (Main Thm)

For cooling schedule $(T_i)_{i \geq 0}$. satisfies:

i) $T_{n+1} < T_n$. $\forall n \geq 0$. ii) $T_n \to 0$ $(n \to \infty)$

iii) $\sum_k e^{-rL/T_{rk-1}} = \infty$. Then, we have:

$$\| A^{(n)}(i, \cdot) - \pi^* \| = \sup_{s \in S} | A^{(n)}(i, s) - \pi^*(s) | \xrightarrow{n \to \infty} 0. \ \forall i$$

<u>Pf:</u> It's application of theory of time-inhomogeneous Markov chain. will be proven later.

Rmk: We can schedule proper temperatures

$(T_n) = (\gamma/\log n)$. $\gamma \geq rL$.

(3) Ergodicity of

Time-inhomogenous MC:

Consider a time-inhomogenous Markov Chain (X_n).

From X_n to X_{n+1}, it has different prob. transition

matrix P_n. Denote: $P^{(m,n)} = \prod_m^{n-1} P_k$.

Def: i) (X_n) is strongly ergodic if $\exists$ dist. π^*

s.t. $\lim_{n\to\infty} \sup_{i\in S} \| P^{(m,n)}(i,\cdot) - \pi^* \| = 0$. $\forall m$

ii) (X_n) is weakly ergodic if:

$\lim_{n\to\infty} \sup_{i,j} \| P^{(m,n)}(i,\cdot) - P^{(m,n)}(j,\cdot) \| = 0$. $\forall m$

Rmk: i) "Weak ergodic" is kind of "loss of memory" after a long time. But: it's unnecessary to converge to some dist.

ii) "$\forall m$" is because we don't want some prob. matrix P_i to determine the whole convergence. (like seq converge)

① Ergodic Coefficient:

Denote: $\delta(P) = \sup_{i,j} \| P(i,\cdot) - P(j,\cdot) \|$. for prob matrix P.

<u>Rmk</u>: $\delta(P)=0 \Rightarrow$ Next move doesn't depend on current state. i.e. lose memory totally.

<u>Lemma</u> $\delta(PQ) \le \delta(P)\delta(Q)$. $\forall P, Q$. prob. matrix.

<u>Pf</u>: Set $A=\{k \mid (PQ)(i,k) > (PQ)(j,k)\}$.

$$\forall i,j\in S: \sum_{k\in S}((PQ)_{ik}-(PQ)_{jk})^+ = \sum_{k\in A}(\square-\square)$$

$$= \sum_{k\in A}\sum_{l\in S} P_{il}Q_{lk} - P_{jl}Q_{lk}$$

$$= \sum_{l\in S}(P_{il}-P_{jl})\sum_{k\in A}Q_{lk}$$

$$\le \sum_S (P_{il}-P_{jl})^+ \sup_l \sum_{k\in A} Q_{lk} - (P_{il}-P_{jl})^- \inf_l \sum_A Q_{lk}$$

$$\overset{\sum P_{il}-P_{jl}=0}{=} \sum_{l\in S}(P_{il}-P_{jl})^+\left(\sup_l \sum_A Q_{lk} - \inf_l \sum_A Q_{lk}\right)$$

$$\le \sum_S (P_{il}-P_{jl})^+\,\delta(Q) \le \delta(P)\delta(Q).$$

<u>Lemma</u>. For dist. λ, μ. and prob matrix P. We have:

$$\|(\lambda-\mu)P\| \le \|\lambda-\mu\|\,\delta(P).$$

<u>Pf</u>. $\forall M \subseteq S$. $\left|\sum_M (\lambda-\mu)P(i)\right| = \left|\sum_{i\in M}\sum_{t\in S}(\lambda-\mu)_t P_{ti}\right|$

$$= \left|\sum_S(\lambda-\mu)_t \sum_{i\in M} P_{ti}\right| \quad (\text{By } \sum(\lambda-\mu)_t = 0)$$

$$\le \sum_{t\in S}(\lambda-\mu)_t^+\left(\sup_t \sum_M P_{ti} - \inf_t \sum_M P_{ti}\right)$$

$$\le \sum_S (\lambda-\mu)_t^+\,\delta(P) = \|\lambda-\mu\|\,\delta(P).$$

<u>Lemma</u> If $\exists$ column of Q. prob. matrix. whose entries $\ge r$.

Then: $\delta(a) \le 1-a$.

Pf. $\delta(a) = \sup_{i,j} \sum_S (a_{ik} - a_{jk})^+ = \sup_{i,j} \sum_A (a_{ik} - a_{jk})$

$\le \sup_{i,j} \sum_A a(i,k) - a \le 1-a$

where suppose that column has index i_0.

WLOG. $i_0 \in A = \{a(i,k) > a(j,k)\}$. Otherwise: consider A^c. since $\sum a_{ik} - a_{jk} = 0$.

prop. If $\exists (n_k) \uparrow \infty$. $\sum (1 - \delta(P^{(n_k, n_{k+1})})) = \infty$. Then: (X_n) is weakly ergodic.

Pf. By $1-x \le e^{-x}$. We have: $\prod_m \delta(P^{(n_k, n_{k+1})}) = 0$. $\forall m$.

$\Rightarrow \delta(P^{(n_k, n_L)}) \le \prod_{i \ge k}^{L-1} \delta(P^{(n_i, n_{i+1})}) \to 0$. $(L \to \infty)$

So: $\delta(P^{(m,n)}) \lesssim \delta(P^{(m, n_k)}) \prod_{i \ge k}^{L(n)} \delta(P^{(n_i, n_{i+1})}) \xrightarrow{n \to \infty} 0$

prop. If (X_n) is weakly ergodic. $\exists (\pi_k)_{k \ge 0}$ stationary dist. for $(P_k)_{k \ge 0}$. s.t. $\sum \| \pi_k - \pi_{k+1} \| < \infty$. Then: (X_n) is strongly ergodic. $\pi^* = \lim_n \pi_n$.

Pf. 1) $\sum \| \pi_k - \pi_{k+1} \| < \infty \Rightarrow \lim \pi_n = \pi^*$ exist.

$\| P^{(k,m)}(i, \cdot) - \pi^*(\cdot) \| \le \| P^{(k,m)}(i, \cdot) - \pi_l P^{(l,m)} \|$

$+ \| \pi_l P^{(l,m)} - \pi_m \| + \| \pi_m - \pi^* \| =: A_1 + A_2 + A_3$

2) $A_1 \le \| P^{(k,l)}(i, \cdot) - \pi_l(\cdot) \| \delta(P^{(l,m)})$

$\to 0$. (as $m \to \infty$). $\forall l$.

3') A_2 : $\lambda_l P^{(l,m)} - \lambda_m = \lambda_l p_l \cdot P^{(l+1,m)} - \lambda_m$

$= \lambda_l \cdot P^{(l+1,m)} - \lambda_m$

$= \lambda_{l+1} P^{(l+1,m)} - \lambda_m + (\lambda_l - \lambda_{l+1}) P^{(l+1,m)}$.

$= \cdots = \sum_l^{m-1} (\lambda_k - \lambda_{k+1}) P^{(k+1,m)}$

$\Rightarrow A_2 \lesssim \sum_l^{m-1} \| \lambda_k - \lambda_{k+1} \| \to 0 \ (m \to \infty)$. $(\delta(p) \leq 2)$.

4') $A_3 \to 0 \ (m \to \infty)$. by 1'). directly.

② Proof of Main Thm:

Denote: Node i_0 is called center. if: $\max_{j \in S} \ell(i_0, j)$

$= r = \min_{S_c} \max_{S} \ell(i,j)$. $D = \max_i d(i)$.

1') Show: $\sum_{k \geq 1} (1 - \delta(P^{(kr-r,kr)})) = \infty$. $P_n =: A_{T_n}$.

Pf: For $i \neq j$: $P_{ij}(n) = \min\{1, e^{-(\ell(j)-\ell(i))/T_n}\} / d(i) \geq e^{-L/T_n} / D$.

$P_{i_0,i_0}(n) = 1 - \sum_{j \in N(i)} P_{ij}(n) \xrightarrow{n \to \infty} 1 - \sum_{\substack{j \in N(i) \\ \ell(j) \leq \ell(i)}} 1/d(i)$

$= \sum_{\substack{j \in N(i) \\ \ell(j) > \ell(i)}} 1/d(i) > 0$ by $i_0 \in S_c$.

$\Rightarrow P_{i_0,i_0}(n) \geq e^{-L/T_n}/D$. for $\forall$ large n.

So: $P^{(m-r,m)}(i, i_0) \geq e^{-rL/T_n} / D^{-r}$. $\forall i \in S$. $\forall m$.

By Lemma: $(1 - \delta(P^{(m-r,m)})) \geq n$.

2') Show: $\sum \| \lambda_n - \lambda_{n+1} \| < \infty$. where $\lambda_n(i) = \dfrac{d(i) e^{-\ell(i)/T_n}}{G(T_n)}$

for strongly ergodic.

Lemma: i) $i \in S^* \Rightarrow \pi_n(i) < \pi_{n+1}(i)$. $\forall n$.

ii) $i \notin S^* \Rightarrow \exists \tilde{n}_i$. s.t. $\pi_{n+1}(i) < \pi_n(i)$. $\forall n \geq \tilde{n}_i$.

Pf: i) is directly check.

ii) $f_i(T) =: e^{-c(i)/T} / G(T)$.

$$f_i'(T) = \lambda(T) \left(\sum_S d(k) e^{-c(k)/T} (c(i) - c(k)) \right)$$

$$\geq \lambda(T) \left(\lambda_1 e^{-\frac{c(i)}{T}} + \lambda_2 e^{-\frac{c(m)}{T}} \right).$$

where $\lambda(T) \geq 0$. $\lambda_1 = \sum_{c(k) \geq c(i)} d(k)(c(i) - c(k))$

and $\lambda_2 = \sum_{c(k) < c(i)} d(k)(c(i) - c(k))$. $c(m) = \min_{\substack{k \\ c(k) < c(i)}} c(k)$

by $i \notin S^*$. ($\{k \mid c(k) < c(i)\} \neq \emptyset$.)

$\Rightarrow$ for sufficient small T. ($T_n \downarrow$ as $n \uparrow$)

Return to pf:

Choose $\hat{n} = \max_{i \notin S^*} \{\tilde{n}(i)\}$. $\sum_{\hat{n}}^{\infty} \| \pi_n - \pi_{n+1} \| = \sum_{\hat{n}}^{\infty} \sum_{i \in S} (\pi_{n+1} - \pi_n)_i^+$

$= \sum_{i \in S^*} \sum_{\hat{n}} (\pi_{n+1}(i) - \pi_n(i))^+$. by monotone Lemma.

$$\Rightarrow \sum_{\hat{n}}^{\infty} \| \pi_n - \pi_{n+1} \| \leq \sum_{i \in S^*} \sum_{\hat{n}} (\pi_{n+1}(i) - \pi_n(i))$$

$$= \sum_{S^*} (\pi^*(i) - \pi_{\hat{n}}(i)) \leq \pi^*(S^*) = 1$$

(4) Card Shuffling:

Q: How close is the desk of cards being random after n shufflings?

i) Rule: Take the top card and insert it into a random position of the desk. equally likely. (May be top again)

ii) Markov Chain:

State space: S_{52} permutation of desk of cards. so that $|S_{52}| = 52!$

initial dist.: $\lambda_0(\beta) = 1$. $\beta = \{$ card k on position $k\}$

($\downarrow$ position 1 → position 52)

Rmk: Stationary λ is: $\lambda(\alpha) = 1/52!$. $\forall \alpha \in S_{52}$.

and note: (X_n) is irred. aperiodic. MC.

$\Rightarrow \|\lambda_n - \lambda\| \to 0$. $(n \to \infty)$

① Cost of random time:

We will find a random time T. s.t. $X_T \sim Uni(52!)$.

Pf: i) U_n is the position of top card move at the n^{th} shuffling. $U_n \overset{i.i.d}{\sim} Unif(52)$.

ii) $T_1 = \inf\{n \geq 1 \mid U_n = 52\}$. $T_k = \inf\{n > T_k \mid U_n \geq 53-k\}$.

Rmk: i) T_k means the time when the top card is inserted below card 52 (so it's in a random position). At time $T = T_{51}+1$. We insert card 52 to a random position. Then we have a desk shuffled randomly!

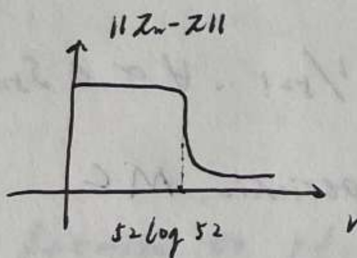
ii) We don't know T_k exactly happen, but know $T = T_{51}+1$. since the top card is 52 at time T_{51}.

For finding $E(T)$:

$T_1 \sim Geo(1/52)$. $T_2 - T_1 \sim Geo(\frac{2}{52}) \ldots T_k - T_{k-1} \sim Geo(\frac{k}{52})$

$T = \sum_2^{52} (T_k - T_{k-1}) + T_1 \Rightarrow E(T) = 52 (\sum_1^{52} 1/k) \sim 52 \log 52.$

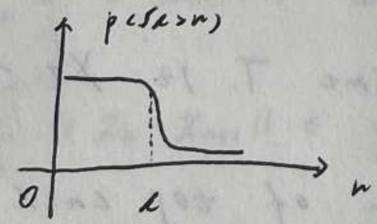
② Threshold Pheonomenon:



$n \uparrow$. $\|Z_n - Z\|$ will have a abrupt decrease from nearly 1 to nearly 0 at $n = [52\log 52]$.

Rmk: Other example:

$X_n \overset{i.i.d}{\sim} N(1,1)$. $S_d = \sum_1^d X_n$. Then near $n=d$:



$p(S_d > n)$ will decrease at order $O(\sqrt{d})$. if d is large enough.

Denote: $A(n) = \|Z_n - Z\|$.

Def: Random time T is strong stationary time if:

i) T is stopping time

ii) X_T indept with T.

iii) $X_T \sim \pi$. stationary dist.

Lemma. If T is strong stationary time for (X_n).

Then: $\|\pi_n - \pi\| \le p(T \ge n)$. $\forall n$.

Pf. 1') Show: $p(T \le n, X_n \in A) = p(T \le n)\pi(A)$.

$$LHS = \sum_{k=1}^{n} p(T=k, X_n \in A)$$

$$= \sum_k \sum_i p(T=k, X_k = i, X_n \in A)$$

$$= \sum_k \sum_i p^{n-k}(i, A)\, p(T=k, X_T = i)$$

$$= p(T \le n)\, \pi(A).$$

2') $\forall A \subseteq S$. $\pi(A) - \pi_n(A) =$

$$\pi(A) - p(X_n \in A, T \le n) - p(X_n \in A, T > n)$$

$$= \pi(A)\, p(T > n) - p(X_n \in A, T > n)$$

$$\Rightarrow |\pi(A) - \pi_n(A)| \le \max\{\square, \square\} \le p(T \ge n)$$

Thm. $\Delta(d\log d + cd) \le p(T > d\log d + cd) \le e^{-c}$. $\forall c > 0$. $d = 52$.

Pf: $T \sim \mathrm{Geo}(1/d) \oplus \cdots \oplus \mathrm{Geo}(d/d)$

(Famous dist. in Coupons collector Problem)

$$\{T > n\} = \bigcup_1^d \{k^{th} \text{ coupon isn't collected at time } n\}$$

$$\triangleq \bigcup_1^d A_k.$$

$$\Rightarrow p(T > n) \le \sum p(A_k) = \sum_1^d \left(\frac{d-1}{d}\right)^n \le d e^{-n/d}$$

Rmk: Fix c. if d is large enough, then cd is small relative to $d\log d$

So it will have an abrupt decrease near $d\log d$ when c changes a lot.

Thm. $k(d) = d\log d - c_d d$. $(c_d) \uparrow \infty$ as $d \to \infty$.

Then: $\| \pi^{(d)}_{k(d)} - \pi^{(d)} \| \xrightarrow{d\to\infty} 1$. index "$(d)$" means dist. π is relative to desk size d.

Pf: We want to find event $A \subseteq S$. st.

$\pi^{(d)}_{k(d)}(A)$ is large (≈ 1). $\pi^{(d)}(A)$ is small.

Consider: $A_{d,a} = \{$ card $d-a+1 \sim d$ are still in their origin order $\}$

$\pi(A_{d,a}) = 1/a!$ (permutate first a cards)

Note: $A_{d,a} = \{$ card $d-a+1$ isn't at top yet $\}$

$\Rightarrow \pi_k(A_{d,a}) \geq p(U > k)$.

U is number of shufflings required for card $d-a+1$ to rise at top.

$U \sim \mathrm{Geo}(a/d) \oplus \cdots \oplus \mathrm{Geo}(d-1/d)$

$$\begin{cases} E(U) = d(\log d - \log a + O(1)) \\ \mathrm{Var}(U) \leq d^2 (\frac{1}{a^2} + \cdots) = O(\frac{1}{a}) d^2 \quad (\mathrm{Var}(\mathrm{Geo}(p)) = \frac{1-p}{p^2}) \end{cases}$$

So: $p(U > k(d)) = p(U - E(U) > -d(c_d - \log a + O(1)))$

$\geq 1 - \mathrm{Var}(U) / d^2 (\square)^2$

choose $a_d = e^{c_d/2}$. ($a = a_d$. depend on d)

Rmk: By the two Thms. above. We know:

$(\overset{d\log d}{\cdot})$ a subtle interval of $d\log d$

contains a abrupt decrease!