6, Stacionární stavy a limitní chování

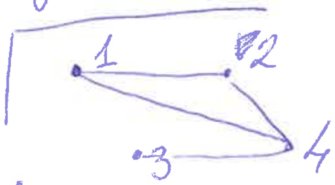
- Stacionární stav Markovova řetězce s maticí pravděpodobností přechodu P je každý stav π, který splňuje
  [stacionární rozdělení]

$$\pi = \pi P.$$

- Zajímá nás existence? jednoznačnost?

  ... Příklady
  - jednoduchá procházka s $p=q=\frac{1}{2}$ na $\mathbb{Z}$ nemá stacionární stav (Cvičení)
  - Stacionární stav nemusí být jednoznačný
    ... pro $P=I$ je každý stav stacionární

- Mějme náhodnou procházku na konečném grafu, tedy $G=(V,E)$, kde $V=\{1,\dots,N\}$ jsou vrcholy a $E\subset\binom{V}{2}=\{(i,j): 1\leq i,j\leq N,\ i<j\}$



$V=\{1,2,3,4\}$
$E=\{(1,2);(1,4);(2,4);(3,4)\}$

Pro $1\leq i\leq N$ označme $\deg(i)$ stupeň vrcholu i
$=\#\{j:\ i \text{ a } j \text{ jsou spojeny hranou}\}$

Pak $\left(\dfrac{\deg(i)}{2\#E}\right)_{i=1}^{N}$ je stacionární stav.

Pozn. $\mathbb{P}(X_{n+1}=j\mid X_n=i)=\begin{cases}0 & \text{pokud } i \,\&\, j \text{ nejsou spojeny hranou}\\ \dfrac{1}{\deg(i)} & \text{pokud } i \,\&\, j \text{ jsou spojeny hranou}\end{cases}$

Skutečně

$$(\pi P)_j = \sum_{i=1}^{N} \pi_i P_{ij} = \sum_{i=1}^{N} \pi_i \mathbb{P}(X_{n+1}=j \mid X_n=i) = \sum_{i=1}^{N} \frac{\deg(i)}{2\#E} \cdot \frac{1}{\deg(i)} \cdot \left(\chi_{(i,j)\in E} + \chi_{(j,i)\in E}\right)$$

$$= \frac{1}{2\#E} \cdot \sum_{j\sim i} 1 = \frac{\deg(j)}{2\#E} = \pi_j \quad \ldots \text{ tedy } \pi = \pi P$$

a navíc $\displaystyle\sum_{j=1}^{N} \pi_j = \frac{1}{2\#E} \sum_{j=1}^{N} \deg(j) = 1.$

- Rovnice detailní rovnováhy
  - Řekneme, že stav π splňuje rovnice detailní rovnováhy, pokud pro každé $i,j \in S$ platí

$$\pi_i P_{ij} = \pi_j P_{ji}.$$

  - Pokud stav π splňuje rovnice detailní rovnováhy, pak je stacionární:

$$(\pi P)_j = \sum_{i\in S} \pi_i P_{ij} = \sum_{i\in S} \pi_j P_{ji} = \pi_j \sum_{i\in S} P_{ji} = \pi_j.$$

Věta: Nechť $(X_n)_{n=0}^{\infty}$ je nerozložitelný Markovský řetězec.

a, Pokud má řetězec stacionární rozložení π, pak nezbytně

π je dáno
$$\pi(i) = \frac{1}{\mu_i}, \quad i \in S,$$

kde $\mu_i = \mathbb{E}[T_i^n | X_0 = i]$ je střední doba návratu $i \in S$.

Tedy π je určeno jednoznačně. Navíc, v tomto případě, jsou všechny stavy pozitivně rekurentní.

b, Naopak, pokud je nerozložitelný řetězec poz. rekurentní, pak stav π na S definovaný $\pi(i) = \frac{1}{\mu_i}$, $i \in S$, je jediný stac. stav.

Důkaz: a, • Nechť π je stac. rozdělení. Řetězec je nerozložitelný, tedy všechny stavy jsou buď rekurentní nebo tranzitní.

Pokud je tranzitní, pak $\lim_{n \to +\infty} P_{ij}^n = 0$ pro všechny $i, j \in S$.

Tedy platí

$$\pi(j) = (\pi P)_j = (\pi P^n)_j = \lim_{n\to\infty} (\pi P^n)_j = \lim_{n\to\infty} \sum_{i\in S} \pi(i) P_{ij}^n$$

$$= \sum_{i\in S} \pi(i) \lim_{n\to\infty} P_{ij}^n = 0, \text{ pro všechna } j \in S$$

Lebesgueova věta

... => π není rozdělení
=> řetězec je rekurentní.

- $$\mu_i = \mathbb{E}[T_i^r \mid X_0 = i] = \sum_{n=1}^{\infty} n \cdot \mathbb{P}(T_i^r = n \mid X_0 = i)$$

$$= \sum_{n=1}^{\infty} \mathbb{P}(T_i^r \geq n \mid X_0 = i) = \sum_{n=1}^{\infty} \frac{\mathbb{P}(T_i^r \geq n \,\&\, X_0 = i)}{\mathbb{P}(X_0 = i)}$$

- pro $n=1$: $\mathbb{P}(T_i^r \geq 1 \,\&\, X_0 = i) = \mathbb{P}(X_0 = i)$ $\boxed{\ldots = \pi(i)!}$
- pro $n \geq 2$ použijeme $\mathbb{P}(A \cap B) = \mathbb{P}(A) - \mathbb{P}(A \cap B^c)$

$$A = \{X_m \neq i, m = 1, \ldots, n-1\}\ ;\ B = \{X_0 = i\}$$

$$\Rightarrow \mathbb{P}(T_i \geq n \,\&\, X_0 = i) = \mathbb{P}(X_1 \neq i, \ldots, X_{n-1} \neq i, X_0 = i)$$

$$= \mathbb{P}(X_1 \neq i, \ldots, X_{n-1} \neq i) - \mathbb{P}(X_0 \neq i, X_1 \neq i, \ldots, X_{n-1} \neq i)$$

$$= \mathbb{P}(X_0 \neq i, \ldots, X_{n-2} \neq i) - \mathbb{P}(X_0 \neq i, X_1 \neq i, \ldots, X_{n-1} \neq i)$$

$$= \alpha_{n-2} - \alpha_{n-1}, \text{ kde } \alpha_n = \mathbb{P}(X_m \neq i \text{ pro } m = 0, 1, \ldots, n) = \mathbb{P}(A_n)$$

$$A_0 \supset A_1 \supset A_2 \supset \ldots \Rightarrow \lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}\Big(\bigcap_n A_n\Big) = \mathbb{P}(X_m \neq i, i \geq 0) = 0$$

(i je rekurentní).

Celkem je tedy $$\mu_i = \sum_{n\geq 1}^{\infty} \frac{\mathbb{P}(T_i^r \geq n \,\&\, X_0 = i)}{\pi(i)} = \sum_{n \geq 2}^{\infty} \frac{\alpha_{n-2} - \alpha_{n-1}}{\pi(i)} + \frac{\pi(i)}{\pi(i)}$$

$$\Rightarrow \mu_i \cdot \pi(i) = \pi_i + \sum_{n=2}^{\infty} (\alpha_{n-2} - \alpha_{n-1}) = \pi_i + \alpha_0 - \lim_{n\to\infty} \alpha_n$$

$$= \mathbb{P}(X_0 = i) + \mathbb{P}(X_0 \neq i) = 1$$

a tedy $\mu_i = \dfrac{1}{\pi(i)}$, resp. $\pi(i) = \dfrac{1}{\mu_i}$.

- Ještě ukážeme, že $\pi(i)>0$ pro každé $i\in S$

Pokud

$$0=\pi(i)=\sum_{k\in S}\pi(k)P^n_{ik}\geq\pi(j)P^n_{ij}\quad\text{pro } \forall n\in\mathbb{N}$$

tak zvolíme n s $P^n_{ij}>0$ & $\pi(j)=0$ ... $\pi\equiv 0$ ... spor.

b, Nechť $(X_n)_{n\geq 0}$ je nerozložitelný a poz. rekurentní.

Tedy $0<\mu_i<\infty$ pro $\forall i\in S$. Definujme

$$\pi(i)=\frac{1}{\mu_i}.$$

Ukážeme, že π je stacionární, tedy
- $\sum_{j\in S}\pi(j)=1$
- $\pi(j)=\sum_{i\in S}\pi(i)P_{ij}$.

Označme $N_j(n)$ počet výskytů v $j\in S$ během prvních n přechodů,

tedy $N_j(n)=\#\{1\leq l\leq n: X_l=j\}=\sum_{l=1}^{n}\chi_{\{X_l=j\}}$.

Pak $\mathbb{E}(N_j(n)|X_0=i)=\sum_{l=1}^{n}\mathbb{P}(X_l=j|X_0=i)=\sum_{l=1}^{n}P^l_{ij}$

a $\sum_{j\in S}\frac{1}{n}\mathbb{E}(N_j(n)|X_0=i)=\frac{1}{n}\sum_{j\in S}\sum_{l=1}^{n}P^l_{ij}=\frac{1}{n}\sum_{l=1}^{n}\underbrace{\sum_{j\in S}P^l_{ij}}_{=1}=1$ ... proto $\forall n$.

Ukážeme, že (*) $\lim_{n\to\infty}\frac{1}{n}\mathbb{E}(N_j(n)|X_0=i)=\frac{1}{\mu_j}\quad\forall i,j\in S$.

Pak $\sum_{j\in S}\pi(j)=\sum_{j\in S}\frac{1}{\mu_j}=\sum_{j\in S}\lim_{n\to\infty}\frac{1}{n}\mathbb{E}(N_j(n)|X_0=i)$

$=\lim_{n\to\infty}\sum_{j}\frac{1}{n}\mathbb{E}(N_j(n)|X_0=i)=1.$

S konečná, jinak technictější

Zbývá dokázat (*): Víme, že $j \in S$ pevné.

Označme T_m čas m-té návštěvy v j, tedy

$$T_m = \min\{n \geq 1 : N_j(n) = m\}. \qquad \dots T_{N_j(n)} \leq n$$

Náhodné proměnné $T_2 - T_1, T_3 - T_2, \dots, T_m - T_{m-1}$ jsou nezávislé a stejně rozdělené.

Tedy
$$\frac{T_m}{m} = \frac{T_1 + (T_2 - T_1) + \dots + (T_m - T_{m-1})}{m} \to \mu_j \qquad = E(T_{k+1} - T_k)$$

Pak pro $T_{N_j(n)} \leq n \leq T_{N_j(n)+1}$ je

$$\frac{T_{N_j(n)}}{N_j(n)} \leq \frac{n}{N_j(n)} \leq \frac{T_{N_j(n)+1}}{N_j(n)}$$

$N_j(n) \to +\infty$ pro $n \to \infty$ (j je poz. rekurentní).

Tedy $\frac{n}{N_j(n)} \to \mu_j$ a $\frac{N_j(n)}{n} \to \frac{1}{\mu_j}$ skoro jistě ... (*) plyne pro podm. střední hodnotu...

- $\pi(j) = \frac{1}{\mu_j}$ je stacionární (pro S konečnou):

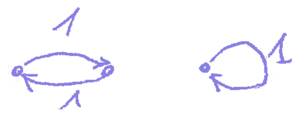
Platí:
$$\sum_{j \in S} \frac{1}{n} E(N_j(n) | X_0 = i) P_{jk} = \sum_{j \in S} \frac{1}{n} \left(\sum_{m=1}^{n} P_{ij}^m \right) P_{jk}$$

$$= \sum_{m=1}^{n} \frac{1}{n} \sum_{j \in S} P_{ij}^m P_{jk} = \sum_{m=1}^{n} \frac{1}{n} P_{ik}^{m+1}$$

$$= \frac{1}{n} \sum_{m=2}^{n+1} P_{ik}^m = \frac{1}{n} \left(\sum_{m=1}^{n+1} P_{ik}^m - P_{ik} \right) = \frac{1}{n} \left\{ E(N_k(n+1) | X_0 = i) - P_{ik} \right\}$$

$n \to \infty$:
$$\sum_{j \in S} \frac{1}{\mu_j} P_{jk} = \frac{1}{\mu_k}. \qquad \boxed{S \text{ konečná, jinak techničtější}}$$

Poznámky: • Pro rozložitelné řetězce může existovat více stac. rozdělení:



má stac. stavy $[\alpha, \alpha, 1-2\alpha]$, $0 \leq \alpha \leq 1/2$.

Otázka: Kdy dojde k tomu, že se rozdělení řetězce limitně blíží nějakému rozdělení, bez ohledu na inicializaci?

Definice (Limitní distribuce): Řekneme, že $(X_n)_{n=0}^{\infty}$ má limitní distribuci, pokud

$$\lim_{n \to \infty} \mathbb{P}(X_n = j \mid X_0 = i) = \lim_{n \to \infty} P^n_{ij}$$

existuje pro každé $i, j \in S$ a tyto limity tvoří rozdělení na S:

$$\forall i \in S: \sum_{j \in S} \lim_{n \to \infty} \mathbb{P}(X_n = j \mid X_0 = i) = 1.$$

V následujícím budeme předpokládat, že S je konečná.

Věta: Pokud pro nějaké $i \in S$ existují $\pi(j) = \lim_{n \to \infty} P^n_{ij}$ pro všechna $j \in S$,

pak π je stacionární rozdělení.

Důkaz: • $\sum_{j \in S} \pi(j) = \sum_{j \in S} \lim_{n \to \infty} P^n_{ij} = \lim_{n \to \infty} \sum_{j \in S} P^n_{ij} = 1$

a

• $\pi(j) = \lim_{n \to \infty} P^{(n+1)}_{ij} = \lim_{n \to \infty} \sum_{k \in S} P^n_{ik} P_{kj} =$

$$\sum_{k \in S} \lim_{n \to \infty} P^n_{ik} P_{kj} = \sum_{k \in S} \pi(k) P_{kj} = [\pi P]_j .$$

Věta: Nechť $(X_n)_{n=0}^{\infty}$ je nerozložitelný a aperiodický a předpokládejme, že $(X_n)_{n=0}^{\infty}$ má stacionární rozdělení π. Nechť $\mathbb{P}(X_0 = j) = \lambda(j)$ je libovolné rozdělení. Pak

$$\mathbb{P}(X_n = j) \to \pi_j \quad \text{pro } n \to \infty \text{ a pro každé } j.$$

Speciálně $P_{ij}^n \to \pi_j$.

Důkaz - metodou couplingu

- Uvažujeme řetězec $(Y_n)_{n=0}^{\infty}$ s poč. stavem $\mathbb{P}(Y_0 = j) = \pi_j$ a maticí přechodu P (tedy stejnou jako $(X_n)_{n=0}^{\infty}$).
- Vezměme $b \in S$ pevné a položme $T = \inf\{n \geq 1 : X_n = Y_n = b\}$
- Ukažme, že $\mathbb{P}(T < +\infty) = 1$: Proces $W_n = (X_n, Y_n)$ je Markovův řetězec na $S \times S$ s pravděpodobností přechodu

$$\tilde{P}_{(i,k)(j,l)} = P_{ij} P_{kl}$$

a počátečním rozdělením $\mu_{(i,k)} = \lambda_i \pi_k$.

Protože X a Y jsou aperiodické, je

$$\tilde{P}^n_{(i,k)(j,l)} = P_{ij}^n P_{kl}^n > 0 \quad \text{pro } n \text{ dostatečně velké.}$$

Také W (s maticí $\tilde{P}$) je tedy aperiodické. W má také stac. distribuci $\tilde{\pi}_{(i,k)} = \pi_i \pi_k$. Tedy je W pozitivně rekurentní. T je čas prvního průchodu W_n bodem (b,b) a je tedy $\mathbb{P}(T < +\infty) = 1$.

• Položme
$$Z_n = \begin{cases} X_n : & n < T \\ Y_n : & n \ge T \end{cases}$$

... Ve chvíli, kdy se X a Y protnou v b, přejde Z z X na Y.

Řetězec $(X_{T+n}, Y_{T+n})_{n \ge 0}^{\infty}$ je Markovský řetězec s maticí $\tilde{P}$ a poč. rozdělením $\delta_{(b,b)}$; nezávislý na $(X_0, Y_0), \dots, (X_T, Y_T)$.

Ze symetrie, uvažujeme-li $(Y_{T+n}, X_{T+n})_{n=0}^{\infty}$, je toto také Markovský řetězec s $\tilde{P}$ a poč. rozdělením $\delta_{(b,b)}$, nez. na $(X_0, Y_0), \dots, (X_T, Y_T)$.

Tedy $W_n' = (Z_n, Z_n')_{n \ge 0}$ je Markovský na $S \times S$, kde
$$Z_n' = \begin{cases} Y_n : & n < T, \\ X_n : & n \ge T, \end{cases}$$
s maticí $\tilde{P}$ a poč. stavem μ.

Tedy $(Z_n)_{n \ge 0}$ je Markovský s maticí P a poč. stavem λ.

• $$\mathbb{P}(Z_n = j) = \mathbb{P}(Z_n = j \;\&\; n < T) + \mathbb{P}(Z_n = j \;\&\; n \ge T)$$
$$= \mathbb{P}(X_n = j \;\&\; n < T) + \mathbb{P}(Y_n = j \;\&\; n \ge T)$$

Tedy $|\mathbb{P}(X_n = j) - \pi_j| = |\mathbb{P}(Z_n = j) - \mathbb{P}(Y_n = j)|$

• X a Z mají stejný poč. stav a stejnou matici P.

• Y má π jako stac. rozdělení ... $\mathbb{P}(Y_n = j) = \mathbb{P}(Y_0 = j) = \pi(j)$

$$= |\mathbb{P}(X_n = j \;\&\; n < T) + \mathbb{P}(Y_n = j \;\&\; n \ge T) - \mathbb{P}(Y_n = j)|$$
$$= |\mathbb{P}(X_n = j \;\&\; n < T) - \mathbb{P}(Y_n = j \;\&\; n < T)| \le \mathbb{P}(n < T) \xrightarrow[n \to \infty]{} 0.$$

Algebraické metody v teorii Markovových procesů

Markovský řetězec je popsán maticí P a poč. stavem X_0. Pokud $\# S = n$, tak $P \in \mathbb{R}^{n \times n}$ a mnohé vlastnosti lze zjistit z vlastností P.

Pokud $A \in \mathbb{R}^{n \times n}$ a $Ax = \lambda x$ ... x je pravý vlastní vektor A

$x^T A = \lambda x^T$ ... x je levý vlastní vektor A

Pro P: součet každého řádku je 1 tedy $P \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$

$\Rightarrow \lambda = 1$ je vlastní číslo & $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ je (pravý) vlastní vektor.

$\Rightarrow \det(A - \lambda I) = 0$ pro $\lambda = 1$ ... existuje i levý vlastní vektor $\pi \neq 0$

... $\boxed{\pi = \pi P}$

... abychom ale mohli π označit za stacionární rozdělení, potřebovali bychom vědět, že $\pi(i) \geq 0$ $\forall i$.

Pak $\frac{1}{\sum_{j=1}^{n} \pi(j)} (\pi(1), \dots, \pi(n))$ je opravdu stac. rozdělení.

Toje obsahem tzv. Perron-Frobeniovy věty

Věta: Nechť $A \in \mathbb{R}^{m\times m}$ je matice s $A_{ij} > 0$ pro $i,j \leq m$. Označme

$\rho = \max_j |\lambda_j|$ její spektrální poloměr. Pak $\rho > 0$ a platí

1, ρ je vlastní číslo A

2, ρ má algebraickou násobnost 1 ... tedy $\det(\lambda I - A) = 0$ má jednoduchý kořen v $\lambda = \rho$.

3, Existuje vlastní vektor v příslušející ρ, který má všechny souřadnice nezáporné (kladné)

4, Pokud $\lambda \neq \rho$ je vlastní číslo A, pak $|\lambda| < \rho$

5, Pokud u je vl. vektor A s kladnými souřadnicemi, pak u je násobek v.

Provedeme alespoň část důkazu

- Nechť $A_{ij} > 0$ • Označme $S = \{x \in \mathbb{R}^m : \|x\|_2 = 1 \ \& \ x_j \geq 0 \text{ pro } 1 \leq j \leq m\}$

  ... je tedy kompaktní

- Pro $x \in S$ je Ax vektor s poz. složkami

- Definujeme $L(x) = \min \left\{ \frac{(Ax)_i}{x_i}, x_i \neq 0 \right\}$

  ... ~~nejmenší~~ největší α s $\forall i \quad \alpha x_i \leq (Ax)_i$

- L je spojitá na S a S je kompaktní => existuje maximální hodnota L na S: $L(v) = \alpha$

• Ukážeme, že α je vl. číslo, v je příslušný vlastní vektor a všechny jeho složky jsou ostře kladné.

• $L(v)=\alpha \Rightarrow \alpha v_i \leq (Av)_i$, tedy $Av-\alpha v \geq 0$ ↑ po složkách

Pokud $Av \neq \alpha v$, tak $A(Av-\alpha v)>0$ a ex. $\varepsilon>0$ s

$$A(Av-\alpha v) > \varepsilon Av$$

Celkem je $\left[A\left(\frac{Av}{\|Av\|_2}\right)\right]_i = \frac{1}{\|Av\|_2}(A(Av))_i > \frac{(\alpha+\varepsilon)}{\|Av\|_2}(Av)_i = (\alpha+\varepsilon)\left(\frac{Av}{\|Av\|_2}\right)_i$

... tedy $L\left(\frac{Av}{\|Av\|_2}\right) \geq (\alpha+\varepsilon)$ ... spor s max. hodnotou L v v.

• $v \geq 0$ protože $v \in S \Rightarrow Av>0$ ... $v>0$.

Dál ukážeme, že α je spektrální poloměr A ... tedy, že $\alpha=\varrho$.

• Nechť $\mu \in \mathbb{C}$ je vl. číslo A a $y \in \mathbb{C}^m$ příslušný vl. vektor s $\|y\|_2=1$.

$1 \leq i \leq m$: $\quad (\mu y)_i = (Ay)_i = \sum_{j=1}^{m} A_{ij} y_j$

$$\Rightarrow |\mu| \cdot |y_i| \leq \sum_{j=1}^{m} A_{ij} \cdot |y_j|$$

Pro vektor $z_i = |y_i|$ platí $\quad |\mu| \cdot z_i \leq (Az)_i$ ... $L(z) \geq |\mu|$

& $L(z) \leq \alpha$ ... $\alpha \geq |\mu|$

... protože α je i vlastní číslo $\Rightarrow \alpha = \varrho$.

$\Rightarrow$ 1, a 3, věty.

... zbytek důkazu vynecháme ...

... Věta platí i pro matice A, pro které ex. $k \in \mathbb{N}$ s $A^k>0$.

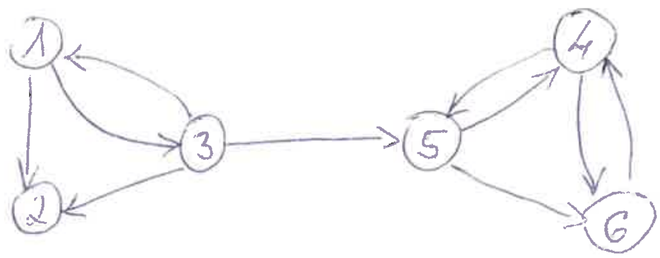
Sergey Brin & Lawrence Page

The Anatomy of a Large-Scale Hypertextual Web Search Engine (1998).

„In this paper, we present Google…"

World Wide Web (WWW) budeme reprezentovat pomocí orientovaného grafu („webgraph"). Uzly jsou stránky, hrany reprezentují odkazy z jedné stránky na druhou.

… Např.:



Algoritmus PageRank je založen na předpokladu, že důležité stránky odkazují na důležité stránky.

„Náhodný klikač" (= „random surfer") startuje z náhodné stránky WWW, a poté kliká na nějakou ze stránek ~~od~~ odkazovaných. Jde tedy o Markovský proces s maticí přechodu pravděpodobností

$$P = \begin{pmatrix} 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ ? & ? & ? & ? & ? & ? \\ 1/3 & 1/3 & 0 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

jak zahrnout stránku, ze které nevede žádný odkaz?

a, $(0\ 1\ 0\ 0\ 0\ 0)$

… byl by absorbující stav

b, $(1/6\ 1/6\ 1/6\ 1/6\ 1/6\ 1/6)$

… náhodný klikač začne opět na nějaké náhodné stránce

Důležitost stránky bude úměrná podílu času, který na ní "náhodný klikač" stráví ... v limitě...

Najdeme tedy stacionární stav tohoto řetězce: $\pi^T = \pi^T P$

V příkladu výše nevede žádný odkaz z $\{4,5,6\}$ do $\{1,2,3\}$...

Stacionární stav by byl tedy koncentrován na $\{4,5,6\}$...

odčerpávají důležitost z celku... Klasik $\{4,5,6\}$ je "absorbující".

Další úprava ... s pravděpodobností α se klikač restartuje...

Nahradíme P pomocí $P' = \alpha P + \frac{1-\alpha}{n} e e^T$ ... $e = (1, \dots, 1)^T$

V našem příkladu $\alpha = 0.9$

$$P' = \begin{pmatrix} 1/60 & 7/15 & 7/15 & 1/60 & 1/60 & 1/60 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 19/60 & 19/60 & 1/6 & 1/6 & 19/60 & 1/6 \\ & & & \ddots & & \end{pmatrix}$$

Důležitost stránky tedy bude dána souřadnicemi stacionárního rozdělení řetězce odpovídajícího P'.

Z <u>Perron-Frobeniovy věty</u> pak plyne, že tento stav je právě jeden, s nezápornými hodnotami...

Navíc všechna ostatní vlastní čísla splňují $|\lambda| < 1$.

Hledání stacionárního stavu

- výpočtem $\pi^T(P'-I)=0$ ... výpočetně neúnosné ... $n \sim 10^{10}$.
- iteracemi ... $\pi^T P' = \sum_{j=1}^{m} \lambda_j \langle \pi_j, \pi \rangle \pi_j$ ... $\lambda_1 = 1$, $|\lambda_j| < 1, j \neq 1$

... tedy $\pi^T (P')^n = \sum_{j=1}^{m} \lambda_j^n \langle \pi_j, \pi \rangle \pi_j$ ... $\lambda_j^n \to 0$ pro $n \to \infty$, $j \neq 1$

- Pro "libovolný" počáteční stav stačí vzít $\pi^T = \lim_{n\to\infty} \pi_0^T (P')^n$

.. konvergence závisí na volbě π_0 a $|\lambda_2|$.

- Iterace $\pi^T \to \pi^T P'$ .. jsou výpočetně náročné, ale schůdné (P' je součet "řídké" matice a jednoduché matice).

Markov Chain Monte Carlo (MCMC)

Monte Carlo - metoda výpočtu deterministických veličin náhodnými algoritmy

Příklad: $\Omega \subset [0;1]^d$, zajímá nás $|\Omega|$ ... Lebesgueova míra Ω.

Předpokládejme, že
- $\int_{[0;1]^d} \chi_\Omega(x)dx$ nelze jednoduše a přesně spočítat
- pro bod $x \in [0;1]^d$ lze jednoduše rozhodnout jestli $x \in \Omega$ nebo $x \notin \Omega$.

$$\Rightarrow |\Omega| = \int_\Omega 1 dx = \int_{[0;1]^d} \chi_\Omega(x)dx \text{ nahradíme } \frac{1}{n}\sum_{i=1}^{n} \chi_\Omega(x_i),$$

kde $x_i \in [0;1]^d$ jsou voleny nezávisle a rovnoměrně rozdělení.

.. obecně $\mathbb{E}X$ nahradíme $\frac{1}{n}\sum_{j=1}^{n} X_j$

Jaká je průměrná chyba?!

$$\mathbb{E}\left|\mathbb{E}X - \frac{1}{n}\sum_{j=1}^{n} X_j\right|^2 = \mathbb{E}\left|\frac{1}{n}\sum_{j=1}^{n}(\mathbb{E}X - X_j)\right|^2$$

$$= \frac{1}{n^2}\,\mathbb{E}\sum_{j,k=1}^{n}(\mathbb{E}X - X_j)(\mathbb{E}X - X_k) = \frac{1}{n^2}\sum_{j=1}^{n}\mathbb{E}(\mathbb{E}X - X_j)^2 = \frac{n\sigma^2}{n^2}$$

$\Rightarrow$ Průměrná chyba (v L_2-smyslu) je tedy $\frac{\sigma}{\sqrt{n}}$, kde $\sigma = var(X)$

a n je počet opakování.

Příklady metody Monte Carlo:

- Výpočet π: $[0,1]^2$, $\Omega = \{(x_1,x_2) \in [0,1]^2, x_1^2 + x_2^2 \le 1\}$

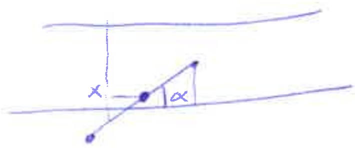
$(x_1,x_2) \in [0,1]^2$ rovnoměrně, $X = \begin{cases} 1 \dots (x_1,x_2) \in \Omega \\ 0 \dots (x_1,x_2) \notin \Omega \end{cases}$

$\mathbb{E}X = \pi/4$ ... Volíme (x_1,x_2) náhodně a počítáme podíl těch, které padnou do jednotkového kruhu...

- Podobně funguje experiment s jehlou délky 1 hozenou náhodně na rovnoběžné přímky se vzdáleností 1:
  - x ... vzdálenost středu jehly od nejbližší přímky ... $[0,1/2]$
  - α – úhel mezi jehlou a přímkami $\alpha \in [0,\pi/2]$

$B = \{(x,\alpha) \in [0,1/2] \times [0,\pi/2] : 0 \le x \le \frac{1}{2}\sin\alpha\}$

$\mathbb{P}(B) = \frac{4}{\pi}\int_0^{\pi/2} \frac{1}{2}\sin\alpha \, d\alpha = \frac{2}{\pi}.$



... Buffon 1777, Laplace 1812

- Vlastnosti náhodných polygonů

Sylvester (1864): Pro konvexní $K \subset \mathbb{R}^2$ zvolme nezávisle $x_1, \dots, x_4 \in K$. Jaká je pravděpodobnost, že $conv(x_1,\dots,x_4)$ je trojúhelník?

... stačí spočítat střední hodnotu plochy $conv(x_1,x_2,x_3)$ ...

$\Rightarrow$ šestirozměrný integrál

... můžeme simulovat metodou Monte Carlo

– nebo spočítat analyticky

(známé hodnoty pro kruh, trojúhelník, ...)

Pokud je integrál tvaru $I = \int_{[0;1]^d} f(x)\psi(x)dx$, pak lze brát

$\frac{1}{n}\sum_{i=1}^{n} f(x_i)$, kde ale x_i jsou brány náhodně s hustotou ψ

- $X = f(x)$, $\mathbb{E}X = \int f(x)\psi(x)dx = I$.

=> Pro použití metody Monte Carlo je nutné umět generovat náhodné vzorky (x_i) vzhledem k dané distribuci.

$\boxed{\text{... Vyrobíme Markovův řetězec, který bude mít danou distribuci jako stacionární rozdělení a vezmeme } X_n \text{ pro } n \text{ velké!}}$

=> Pro S a π na S dané hledáme Markovský řetězec tak, aby

- π byl stacionární stav ... $\pi = \pi P$
- X_n konvergovalo rychle k π (v nějakém smyslu)

  ... tedy aby $\mathbb{P}(X_n = j) \sim \pi_j$ pro n malé...

- Další podmínky ... $X_n \to X_{n+1}$ lze realizovat rychle

Mějme tedy dáno S a π na S.

Metropolisův algoritmus:

- Předpokládejme, že na S máme (nerozložitelný) Markovský řetězec se symetrickou maticí pravděpodobností přechodu $(\Psi_{x,y})_{x,y\in S}$
- Modifikujeme tento řetězec tak, aby měl stac. rozdělení π.
- Přechod z x do y akceptujeme s pravd. $a(x,y)$, odmítneme s $1-a(x,y)$

Tedy
$$P_{xy} = \begin{cases} \Psi_{xy}\, a(x,y), & y\neq x \\ 1-\sum\limits_{y\neq x}' \Psi_{xy}\, a(x,y), & y=x \end{cases}$$

- π bude stac. pro P, pokud budou splněny rovnice detailní rovnováhy:

$$\pi(x) P_{xy} = \pi(y) P_{yx} \quad , x\neq y$$

tedy
$$\pi(x) \cancel{\Psi_{xy}}\, a(x,y) = \pi(y) \cancel{\Psi_{yx}}\, a(y,x)$$

Ψ symetrická

$$\pi(x) a(x,y) = \pi(y) a(y,x)$$

a chceme volit co největší ($\Rightarrow$ rychlost konvergence)

$$\Rightarrow \pi(x) a(x,y) = \pi(y) a(y,x) = \min(\pi(x), \pi(y))$$

$$\ldots a(x,y) = \frac{\min(\pi(x),\pi(y))}{\pi(x)}, \quad a(y,x) = \frac{\min(\pi(x),\pi(y))}{\pi(y)}$$

$$= \min\left(1, \frac{\pi(y)}{\pi(x)}\right)$$

$$\Rightarrow P_{xy} = \begin{cases} \Psi_{xy}\cdot \min\left(1, \frac{\pi(y)}{\pi(x)}\right), & x\neq y \\ 1-\sum\limits_{y\neq x}' \downarrow & , x=y \end{cases}$$