5.1. Markov chains

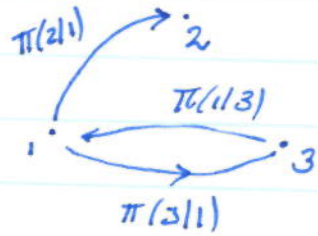
- Sequence of RVs: $X_1, X_2, \dots, X_n$ $\qquad X_i \in \Lambda$ ← discrete
- Markov chain joint pdf:

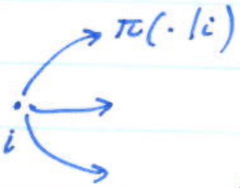
$$p(X_1, X_2, \dots, X_n) = p(X_1)\, \underbrace{\pi(X_2|X_1)}_{\text{correlation - not iid!}}\, \pi(X_3|X_2) \cdots \pi(X_n|X_{n-1})$$

- Transition probability: $\pi(y|x) = Pr(X_n = y \mid X_{n-1} = x) = P(x \to y)$
  - Assumed time-independent (homogeneous)
  - Normalization:

$$\sum_{y \in \Lambda} \pi(y|x) = 1 \quad \forall\, x$$

  - Graphical representation





$\sum \pi(\cdot|i) = 1$

since must jump with prob=1

  - Markov chain representation:

$$X_1 \to X_2 \to X_3 \to \cdots \to X_{n-1} \to X_n$$

↑ time

↳ transition depend on previous RV - neighbor

- Propagation:

$$P(X_n = y) = \sum_{x \in \Lambda} \pi(y|x)\, P(X_{n-1} = x)$$

or

$$p_n(y) = \sum_x \pi(y|x)\, p_{n-1}(x)$$

or

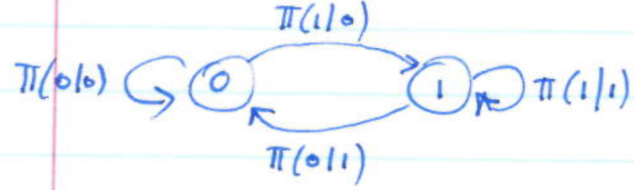
$$\underbrace{p_n}_{\text{final}} = \underbrace{\Pi}_{\text{Matrix}}\, \underbrace{p_{n-1}}_{\text{initial}}$$

vectors

Example: 2 state MC

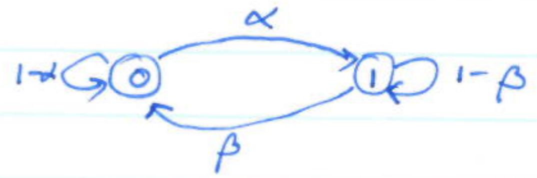
$X_1, X_2, \ldots, X_n \qquad X_i \in \{0,1\}$

(p_0, p_1) start



$p_0' = \pi(0|0)\, p_0 + \pi(0|1) p_1$

$p_1' = \pi(1|0)\, p_0 + \pi(1|1) p_1$



$p_0' = (1-\alpha) p_0 + \beta p_1$

$p_1' = \alpha\, p_0 + (1-\beta) p_1$

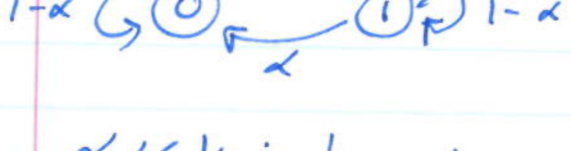
$$\begin{pmatrix} p_0' \\ p_1' \end{pmatrix} = \begin{pmatrix} \pi(0|0) & \pi(0|1) \\ \pi(1|0) & \pi(1|1) \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \end{pmatrix} = \begin{pmatrix} 1-\alpha & \beta \\ \alpha & 1-\beta \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \end{pmatrix}$$

- Columns sum to 1
- $\pi(j|i) \qquad \sum_j \pi(j|i) = 1 \quad \forall\, i$

  j: row, i: colum

$$p' = \Pi p$$
$$p_2 = \Pi p_1$$

Symmetric case



$$\Pi = \begin{pmatrix} 1-\alpha & \alpha \\ \alpha & 1-\alpha \end{pmatrix}$$

$\alpha \ll 1/2$: Long sequences of 0's and 1's $\qquad$ 000..01111.. (rare jump)

$\alpha \gg 1/2$: Long alternating sequences $\qquad$ 010100 (rare persistence)

- Ergodic MCs

$$p_\infty = \lim_{n\to\infty} p_n = \lim_{n\to\infty} \Pi^{n-1} p_1$$

exists and is the same $\forall$ initial pdf p_1

- Equivalent to connected (communicating) MC:

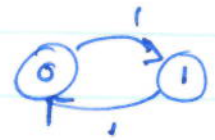
$\forall x, y \ \exists j$ such that

$$P(X_j = y \mid X_i = x) > 0 \qquad j > i$$

$\Rightarrow$ Always possible to reach y from x in finite time $j-i$.

Example: $\Pi = \begin{pmatrix} \alpha-1 & \beta \\ \alpha & 1-\beta \end{pmatrix}$ ergodic for $0<\alpha<1$, $0<\beta<1$

$\Pi = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  not ergodic

$\Pi = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  cyclic not ergodic

5.2. Large deviations of additive Markov functionals (Observables or RVs)

- $X_1, X_2, \dots, X_n$ MC, homogeneous
  Transition matrix $\Pi(y|x)$ $x, y \in \Lambda$

- $S_n = \frac{1}{n} \sum_{i=1}^{n} X_i$

→ LDP for S_n and rate fct ?

- Gärtner-Ellis theorem: $I(s) = \sup_k \{ks - \lambda(k)\}$
- SCGF:
  $$\lambda(k) = \lim_{n \to \infty} \frac{1}{n} \ln E[e^{nkS_n}]$$

- Explicit calculation:
  $$E[e^{nkS_n}] = E[e^{k\Sigma_i X_i}]$$
  $$= \sum_{X_1 X_2 \dots X_n} p(X_1 \dots X_n)\, e^{k\Sigma_i X_i}$$
  $$= \sum_{X_1 \cdots X_n} p(X_1)\, \Pi(X_2|X_1) \cdots \Pi(X_n|X_{n-1})\, e^{k\Sigma_i X_i}$$
  $$= \sum_{X_1 \cdots X_n} p(X_1) e^{kX_1}\, \Pi(X_2|X_1) e^{kX_2} \cdots \Pi(X_n|X_{n-1}) e^{kX_n}$$

$p_k(x) = p(x)e^{kx}$
$\Pi_k(y|x) = \Pi(y|x)e^{ky}$

$$= \sum_{X_1 \cdots X_n} p_k(X_1)\, \Pi_k(X_2|X_1) \cdots \Pi_k(X_n|X_{n-1})$$
$$= \sum_{X_n \cdots X_1} \Pi_k(X_n|X_{n-1})\, \Pi_k(X_{n-1}|X_{n-2}) \cdots \Pi_k(X_2|X_1)\, p_k(X_1)$$
$$= \sum_{X_n} \sum_{X_{n-1}} \Pi_k(X_n|X_{n-1}) \sum_{X_{n-2}} \Pi_k(X_{n-1}|X_{n-2}) \cdots \underbrace{\sum_{X_1} \Pi_k(X_2|X_1)\, p(X_1)}_{\Pi_k p}$$
$$= \sum_{X_n} \left(\Pi_k^{n-1} p_k\right)(X_n)$$
$$= \mathrm{Tr}\left(\Pi_k^{n-1} p_k\right) \leftarrow \text{vector}$$

- Behavior as $n \to \infty$
  - Π is a positive matrix (stochastic matrix)
  - Π_k " " " " (tilted matrix)

  - Perron-Frobenius Theorem:
    - $A > 0$
    - $A v_i = \zeta_i v_i$ eigenvalue eigenvector
    - $\zeta_1 < \zeta_2 < \cdots < \zeta_{max}$

  $\Rightarrow A x \xrightarrow{n\to\infty} \text{const} \cdot \zeta_{max}^n v_{max}$

  - Similar result for $A \geq 0$

- Generating function:

$$\Rightarrow E[e^{nkS_n}] = Tr(\Pi_k^{n-1} p_k)$$

$$\sim \text{const} \cdot \zeta_{max}^{n-1}(\Pi_k)$$

$$= \text{const}\ \zeta_{max}^{n}(\Pi_k)$$

- SCGF: $\Rightarrow \lambda(k) = \ln \zeta(\Pi_k)$

- Calculation steps:

1- Write matrix Π_k : $\Pi_k(\underset{\text{row}}{y}|\underset{\text{column}}{x}) = \Pi(y|x)e^{ky}$

2- Find dominant eigenvalue : $\zeta(\Pi_k)$

3- Get $\lambda(k) = \ln \zeta(\Pi_k)$

4- Check conditions of Gärtner-Ellis Theorem

5- Calculate Legendre transform:

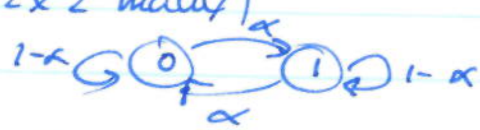
$$I(s) = k_s \cdot s - \lambda(k_s) \qquad k_s : \lambda'(k) = s$$

Co

Example: 2 state MC, symmetric α

$\Lambda = \{0, 1\}$ (2 x 2 matrix)

$$\Pi = \begin{pmatrix} 1-\alpha & \alpha \\ \alpha & 1-\alpha \end{pmatrix}$$



$$S_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$\Pi_k = \begin{pmatrix} (1-\alpha) e^{k0} & \alpha e^{k0} \\ \alpha e^{k\cdot 1} & (1-\alpha) e^{k\cdot 1} \end{pmatrix} = \begin{pmatrix} 1-\alpha & \alpha \\ \alpha e^k & (1-\alpha) e^k \end{pmatrix}$$

$\hookrightarrow$ Get $\rho(\Pi_k)$, $\lambda(k)$, $I(s)$. <u>Exercise</u>.

Generalization 1:

$S_n = \frac{1}{n}\sum_{i=1}^{n} f(X_i)$ instead of $\frac{1}{n}\sum_{i=1}^{n} X_i$

$\Rightarrow$ Π_k : $\Pi_k(y|x) = \pi(y|x)\, e^{k f(y)}$ tilted matrix

Generalization 2:

$$S_n = \frac{1}{n}\sum_{i=1}^{n-1} g(X_i, X_{i+1})$$

$\Pi_k(y|x) = \pi(y|x)\, e^{k g(x,y)}$ <u>Exercise</u>

Example: Same 2 state MC, symmetric α

$S_n = \frac{1}{n}\sum_{i=1}^{n-1} \delta_{X_i, X_{i+1}}$ persistence

$S_n = \frac{1}{n}\sum_{i=1}^{n-1} \left(1 - \delta_{X_i, X_{i+1}}\right)$ mean jump # = current

5.3. Level 2 Markov large deviations

- Markov extension of Sanov Theorem
- $X_1, X_2, \ldots, X_n$ MC, Π, Λ
- Empirical vector density:

$$L_n(x) = \begin{cases} \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i, x} & \text{discrete} \\ \frac{1}{n} \sum_{i=1}^{n} \delta(X_i - x) & \text{continuous} \end{cases}$$

- LDP: $P(L_n = l) \asymp e^{-n I(l)}$
- iid case (Sanov): $I(l) = D(l \| p)$
- Markov case:

$$I(l) = \sup_{u>0} \sum_{i \in \Lambda} l_i \ln \frac{u_i}{(\Pi u)_i}$$

$$= - \inf_{u>0} \sum_{i \in \Lambda} l_i \ln \frac{(\Pi u)_i}{u_i}$$

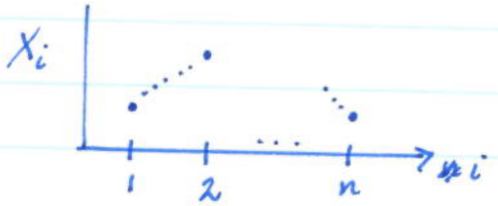
↳ not explicit

Proof: (reading) From Gärtner-Ellis Theorem.
Dembo & Zeitouni, Sec. 3.1.2 p 76 -

* 5.4. Continuous-time processes

Discrete time

$X_1, X_2, \dots, X_n \quad \{X_i\}_{i=1}^{n}$

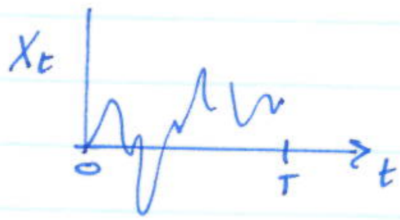


$\Pi: \quad \Pi(y|x) = P(x \to y)$

$S_n = \frac{1}{n} \sum_{i=1}^{n} f(X_i)$

Continuous time

$X_t, \ t \in [0,T] \quad \{X_t\}_{t=0}^{T}$



$$\Pi_{\Delta t}: \ P(X_{t+\Delta t} = y | X_t = x) = \Pi_{\Delta t}(y|x) = \underbrace{w(x \to y)}_{\text{jump rate}} \Delta t$$

$$S_T = \frac{1}{T} \int_0^T f(X_t)\,dt$$

- ~~Time discretization:~~
- $\cancel{\Pi_{\Delta t}(y|x) = w(x \to y) \Delta t = e^{\Delta t L(x \to y)}}$
- Result: $\lambda(k) = \zeta_{\max}(L_k)$

L: generator

L_k = tilted generator

$$= L + \underbrace{k \Pi f(x)}_{\text{diagonal perturbation}}$$

Proof (reading): Sec 3.3 of HT2011