Summary

\{X_n, n \geq 0\} - Markov chain

Countable or finite state space \( S = \{0, 1, 2, \ldots\} \)

\[ \mathbb{P}(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = P_{ij} \]

\( P_{ij} \geq 0, \sum_j P_{ij} = 1 \)

\( \mathbf{P} = (P_{ij}) \) – transition matrix

\( P^n = (P^n_{ij}) \) – \( n \)-step transition probabilities

\( T_i = \inf\{n > 0 : X_n = i\}, i \in S \) – 1st hitting time of state \( i \)

\( m_{ii} = \mathbb{E}(T_i | X_0 = i) \) – mean return time to state \( i \)
Weather example

\[ P(\text{rain tomorrow} \mid \text{rain today}) = 0.7 \]

\[ P(\text{rain tomorrow} \mid \text{no rain today}) = 0.4 \]

\[
\mathbf{P} = \begin{bmatrix}
0.7 & 0.3 \\
0.4 & 0.6
\end{bmatrix} \\
\mathbf{P}^2 = \begin{bmatrix}
0.61 & 0.39 \\
0.52 & 0.48
\end{bmatrix} \\
\mathbf{P}^4 = \begin{bmatrix}
0.5749 & 0.4251 \\
0.5668 & 0.4332
\end{bmatrix} \\
\mathbf{P}^8 = \begin{bmatrix}
0.5715 & 0.4285 \\
0.5714 & 0.4286
\end{bmatrix}
\]

So, after many days, today’s weather does not affect the forecast anymore. The chance of rain will ‘stabilize’

\( \pi_0 \) - chance of rain, \( \pi_1 \) - chance of no rain, after infinitely long time.

Linear equations:

\[
\pi_0 = 0.7\pi_0 + 0.4\pi_1, \quad \pi_1 = 0.3\pi_0 + 0.6\pi_1, \quad \pi_0 + \pi_1 = 1
\]

Solution: \( \pi_0 = 8/14 \approx 0.5714, \quad \pi_1 = 6/14 \approx 0.4286 \)
Convergence theorem

Def. Positive recurrent aperiodic states are called *ergodic*

Theorem. For an irreducible ergodic Markov chain, \( \lim_{n \to \infty} P^n_{ij} \) exists and is independent of \( i \). Furthermore, letting

\[
\pi_j = \lim_{n \to \infty} P^n_{ij}, \quad j \geq 0
\]

then \( \pi_j \) is the unique non-negative solution of

\[
\pi_j = \sum_i \pi_i P_{ij}, \quad j \geq 0, \quad \sum_i \pi_i = 1.
\]

Moreover, \( \pi_j = 1/m_{jj} \) where \( m_{jj} \) is the expected time between two successive visits to \( j \).

\( \pi_j \)’s represent a *stationary* distribution of a Markov chain

Remark: A finite aperiodic irreducible Markov chain is ergodic
Example 4.19/4.22 (A model of class mobility)

Occupation classes: 0: lower; 1: middle; 2: upper. Assume that occupation of children depends only on the occupation of parents, and the transition matrix is:

\[
P = \begin{pmatrix}
0.45 & 0.48 & 0.07 \\
0.05 & 0.70 & 0.25 \\
0.01 & 0.5 & 0.49
\end{pmatrix}
\]

In the long-run, what percentage of population will have lower-, middle-, or upper-class jobs?
Solution. The limiting probabilities $\pi_i$ satisfy

\[
\begin{align*}
\pi_0 &= 0.45\pi_0 + 0.05\pi_1 + 0.01\pi_2 \\
\pi_1 &= 0.48\pi_0 + 0.70\pi_1 + 0.25\pi_2 \\
\pi_2 &= 0.07\pi_0 + 0.25\pi_1 + 0.49\pi_2 \\
\pi_0 + \pi_1 + \pi_2 &= 1
\end{align*}
\]

Hence, $\pi_0 = 0.07, \pi_1 = 0.62, \pi_2 = 0.31$
**Linear algebra point of view**

Consider a finite irreducible MC with \( n \) states

Let vector \( \pi \) be defined as \((\pi_1, \pi_2, \ldots, \pi_n)\)

Then stationary probabilities satisfy \( \pi P = \pi \), so \( \lambda = 1 \) is an eigenvalue and \( \pi \) is the corresponding eigenvector of \( P \). The whole problem of finding \( \pi \) can be formulated in framework of linear algebra!

Applying *Perron-Frobenius theory*: If \( P \) induces an irreducible aperiodic MC then \( \lambda_1 = 1 \) is a simple eigenvalue of \( P \), and all other eigenvalues \( \lambda_2, \ldots, \lambda_p \), where \( p \leq n \), lie on a disk \(|z| < 1|\).

The existence and uniqueness of \( \pi \) can be proved as well as convergence. This is however not a simple theory!
Example 4.21/4.24 (partially)

A production process changes its states according to a positive recurrent Markov chain with transition probabilities $P_{ij}, i, j = 1, \ldots, n$. The process is up in a set $A$ and down in a set $A^c$. What is the average time between two break-downs?

**Solution.** Rate visit $i = \pi_i$

Rate enter $j$ from $i = \pi_i P_{ij}$

Rate enter $j$ from $A = \sum_{i \in A} \pi_i P_{ij}$

Rate breakdown occurs $= \sum_{j \in A^c} \sum_{i \in A} \pi_i P_{ij}$

Average time between breakdowns $= \left[ \sum_{j \in A^c} \sum_{i \in A} \pi_i P_{ij} \right]^{-1}$

(see numerical example in the book)
Google PageRank


PageRank $\pi_i$ of page $i$ is the long run fraction of time that a random surfer spends on page $i$.

‘Easily bored surfer’ model. With probability $c$ (=0.85), a surfer follows a randomly chosen outgoing link. Otherwise, he/she jumps to a random page.

$$\pi_i = \sum_{j \to i} \frac{c}{d_j} \pi_j + \frac{1-c}{n}$$

Page is important if many important pages link to it!

Is the PageRank well defined? Why?
Computational issues

Problem: Find a stationary distribution of a (large) finite MC.

Coupling argument: convergence ‘happens’ at some random time $T$ (mixing time, coupling time). But... this is fairly useless in practice

**Power method.** $\pi^{(0)}$ - arbitrary initial distribution;

$$\pi^{(k)} = \pi^{(k-1)}P = \pi^{(0)}P^k$$

Since $P^k_{ij} \to \pi_j$, we have $\pi^{(k)}_j = \sum_i \pi^{(0)}_i P^k_{ij} \to \pi_j$ as $k \to \infty$

Moreover, one can show that $|\pi^{(k)}_j - \pi_j| = O(|\lambda_2|^k)$, where $1 > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_p|$ are the eigenvalues of $P$.

Thus, we have exponential convergence. Problems arise if $|\lambda_2| \approx 1$. There are many advanced methods that speed up the convergence.

Notes on convergence

Stationary distribution $\pi$: $\sum_i \pi_i P_{ij} = \pi_j$, $\sum_j \pi_j = 1$

- MC chain irreducible and ergodic (all states are positive recurrent and aperiodic) $\Rightarrow \lim_{n \to \infty} P_{ij}^n \to \pi_j = 1/m_{jj} > 0$ for all $i, j \in S$

- $j$ is transient or null-recurrent $\Rightarrow \lim_{n \to \infty} P_{ij}^n = 0$, $i \in S$. All states trans. or null-rec. $\Rightarrow$ stationary distribution does not exist.

- MC is irreducible, positive recurrent but periodic with period $d$ $\Rightarrow \pi_j = 1/m_{jj}$, $j \in S$, is a stationary distribution, but $\lim_{n \to \infty} P_{ij}^n$ does not exists. Interpret $\pi_j$ as the fraction of time spent in $j$ or consider $\lim_{n \to \infty} P_{jj}^n = d/m_{jj}$

- MC is reducible with several positive recurrent classes $\Rightarrow$ stationary distribution is not unique

- It can be shown that $\pi_j$ is the long-run fraction of time spent in $j$ (the Law of Large Numbers)
Strong Law of Large Numbers

**Proposition 4.3** Let \( \{X_n, n \geq 1\} \) be an irreducible Markov chain with stationary probabilities \( \pi_j, j \geq 0 \) and let \( r \) be a bounded function on the state space. Then, w.p. 1,

\[
\lim_{N \to \infty} \frac{\sum_{n=1}^{N} r(X_n)}{N} = \sum_{j=0}^{\infty} r(j) \pi_j
\]

**Proof.** Let \( a_j(N) \) be the number of visits to state \( j \) during the time \( 1, \ldots, N \), then

\[
\lim_{N \to \infty} \frac{\sum_{n=1}^{N} r(X_n)}{N} = \sum_{j=0}^{\infty} r(j) \lim_{N \to \infty} \frac{a_j(N)}{N}.
\]

Then apply the Law of Large Numbers.
Example 4.23 (Average premium in Bonus Malus car insurance)

4 states for the premium. Assume that the number of claims by a customer has a Poisson distribution with mean 1/2. Then the transition matrix is:

\[
P = \begin{pmatrix}
0.6065 & 0.3033 & 0.0758 & 0.0144 \\
0.6065 & 0.0000 & 0.3033 & 0.0902 \\
0.0000 & 0.6065 & 0.0000 & 0.3935 \\
0.0000 & 0.0000 & 0.6065 & 0.3935
\end{pmatrix}
\]

Hence, \( \pi_1 = 0.3692, \pi_2 = 0.2395, \pi_3 = 0.2103, \pi_4 = 0.1809 \)

The function \( r(i) \) is a premium in state \( i \). Thus,
\( r(1) = 200, r(2) = 250, r(3) = 400, r(4) = 600 \).

The long-run average premium per year is
\[
200\pi_1 + 250\pi_2 + 400\pi_3 + 600\pi_4 = 326.375
\]
Some history

St.Petersburg school:
P.L. Chebyshev and his students
A.A. Markov, A.M. Lyapunov, G.F. Voronoi

A.A. Markov (1856-1922)

Weak law of large numbers (J.Bernoulli, 1713): \(X_1, X_2, \ldots\) - i.i.d.

\[
\lim_{n \to \infty} P \left\{ \left| \frac{X_1 + \cdots + X_n}{n} - E(X) \right| \geq \varepsilon \right\} = 0
\]

P.A. Nekrasov (1898, 1902): Law of large numbers holds only for independent random variables. In 1906, A.A. Markov presented a sequence of dependent random variables for which the Bernoulli’s theorem holds. In 1926, S.N. Bernstein called this type of sequences a Markov chain
The Gambler Ruin Problem (section 4.5.1)

A Gambler wins a game w.p. $p$ and loses w.p. $q = 1 - p$. Initial fortune is $i$. What is the chance that the Gambler reaches fortune $N$ before going broke?

$X_n$ - the player’s fortune at time $n$. Clearly, $\{X_n, n \geq 0\}$ – MC

$$P_{00} = P_{NN} = 1; \quad P_{i,i+1} = p = 1 - P_{i,i-1}, \ i = 1, \ldots, N - 1$$

$P_i, i = 0, \ldots, N$ - prob-ty to reach $N$ starting with $i$

$$P_i = pP_{i+1} + qP_{i-1}, \ i = 1, \ldots, N - 1; \quad P_N = 1; \quad P_0 = 0$$

Equivalently, since $p + q = 1$,

$$(p + q)P_i = pP_{i+1} + qP_{i-1}, \ i = 1, \ldots, N - 1$$
The Gambler Ruin Problem (continued)

Denote $Z_i = P_i - P_{i-1}$. Then from the last equation,

$$Z_{i+1} = \frac{q}{p}Z_i, \ i = 1, \ldots, N - 1$$

By induction, we get $Z_i = Z_1 (q/p)^{i-1}, \ i = 1, \ldots, N$. Also,

$$P_i = Z_1 + \cdots + Z_i, \ i = 1, \ldots, N$$

**Case 1.** If $p \neq q$, then $P_N = Z_1 \frac{1 - (q/p)^N}{1 - q/p} = 1$

So, $P_1 = Z_1 = \frac{1 - q/p}{1 - (q/p)^N}$ and $P_i = \frac{1 - (q/p)^i}{1 - (q/p)^N}, \ i = 1, \ldots, N$

**Case 2.** If $p = q$ then $Z_1 = Z_2 = \ldots = Z_N \Rightarrow P_N = NZ_1 = 1.$

So, $Z_1 = 1/N$ and $P_i = i/N, \ i = 1, \ldots, N$
The Gambler Ruin Problem (transient states)

Let $p = 0.4$ and $q = 0.6$, $N = 7$, initial capital=3

$$P_T = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 0 & 0.4 & 0 & 0 & 0 \\ 2 & 0.6 & 0 & 0.4 & 0 & 0 \\ 3 & 0 & 0.6 & 0 & 0.4 & 0 \\ 4 & 0 & 0 & 0.6 & 0 & 0.4 \\ 5 & 0 & 0 & 0 & 0.6 & 0 \\ 6 & 0 & 0 & 0 & 0 & 0.6 \end{bmatrix}$$

$$S = (s_{ij}) = [I - P_T]^{-1}$$

$$\mathbb{E}[\text{#times the gambler has 5 units}] = s_{3,5} = 0.9228$$

$$\mathbb{E}[\text{#times the gambler has 2 units}] = s_{3,2} = 2.3677$$

$$P[\text{gambler ever has a fortune of 1}] = f_{3,1} = s_{3,1}/s_{1,1} = 0.8797$$

**Check:** last prob-ty is equivalent to

$$P(2 \text{ steps down before 4 steps up}) = 1 - \frac{1 - (0.6/0.4)^2}{1 - (0.6/0.4)^6} = 0.8797$$