Proof. (i) Assume $i \in C$, C minimal closed subset of S. Then for any $k \geq 1$,

$$\sum_{j \in S} p_{ij}^{(k)} = \sum_{j \in C} p_{ij}^{(k)} = 1,$$

because C is closed and P is a stochastic matrix. Consequently,

$$\sum_{k \geq 0} \sum_{j \in C} p_{ij}^{(k)} = \infty,$$

and because C is finite, there must be some $j_0 \in C$ such that

$$\sum_{k \geq 0} p_{ij_0}^{(k)} = \infty.$$

Since $j_0 \leftrightarrow i$, there is some $k_0 \geq 0$ such that $p_{j_0 i}^{(k_0)} = p_0 > 0$. But then

$$\sum_{k \geq 0} p_{ii}^{(k)} \geq \sum_{k \geq k_0} p_{ij_0}^{(k-k_0)} p_{j_0 i}^{(k_0)} = \left(\sum_{k \geq k_0} p_{ij_0}^{(k-k_0)} \right) \cdot p_0 = \infty.$$

By Theorem 1.4 i is thus recurrent.

(ii) Denote $C = C_1 \cup \cdots \cup C_m$. Since for any $j \in Y$ the set $\{l \in S \mid j \to l\}$ is closed, it must intersect C; thus for any $j \in T$ there is some $k \geq 1$ such that

$$p_{iC}^{(k)} \triangleq \sum_{l \in C} p_{jl}^{(k)} > 0.$$

Since T is finite, we may find a $k_0 \geq 1$ such that for any $j \in T$, $p_{jC}^{(k_0)} = p > 0$. Then one may easily compute that for any $i \in T$,

$$p_{iT}^{(k_0)} \leq 1 - p, \; p_{iT}^{(2k_0)} \leq (1-p)^2, \; p_{iT}^{(3k_0)} \leq (1-p)^3, \text{ etc.}$$

and so

$$\sum_{k \geq 1} p_{ii}^{(k)} \leq \sum_{k \geq 1} p_{iT}^{(k)} \leq \sum_{r \geq 0} k_0 p_{iT}^{(rk_0)} \leq k_0 \sum_{r \geq 0} (1-p)^r < \infty.$$

By Theorem 1.4, i is thus transient. □

1.2 Existence and Uniqueness of Stationary Distribution

A matrix $A \in \mathbb{R}^{n \times n}$ is

(i) *nonnegative*, denoted $A \geq 0$, if $a_{ij} \geq 0 \quad \forall\, i, j$

(ii) *positive*, denoted $A \gtrsim 0$, if $a_{ij} \geq 0 \quad \forall\, i, j$ and $a_{ij} > 0$ for at least one ij

(iii) *strictly positive*, denoted $A > 0$, if $a_{ij} > 0 \quad \forall\, i, j$

We denote also $A \geq B$ if $A - B \geq 0$, etc.

Lemma 1.6 *Let $P \geq 0$ be the transition matrix of some regular finite Markov chain with state set S. Then for some $t_0 \geq 1$ it is the case that $P^t > 0 \quad \forall\, t \geq t_0$.*

Proof. Choose some $i \in S$ and consider the set

$$N_i = \{t \geq 1 \mid p_{ii}^{(t)} > 0\}.$$

Since the chain is (finite and) aperiodic, there is some finite set of numbers $t_1, \ldots, t_m \in N_i$ such that

$$\gcd N_i = \gcd\{t_1, \ldots, t_m\} = 1,$$

i.e. for some set of coefficients $a_1, \ldots, a_m \in \mathbb{Z}$,

$$a_1 t_1 + a_2 t_2 + \cdots + a_m t_m = 1.$$

Let P and N be the absolute values of the positive and negative parts of this sum, respectively. Thus $P - N = 1$. Let $T \geq N(N-1)$ and consider any $s \geq T$. Then $s = aN + r$, where $0 \leq r \leq N-1$ and, consequently, $a \geq N-1$. But then $s = aN + r(P-N) = (a-r)N + P$ where $a - r \geq 0$, i.e. S can be represented in terms of $t_1, \ldots, t_m$ with nonnegative coefficients $b_1, \ldots, b_m$. Thus

$$p_{ii}^{(s)} \geq p_{ii}^{(b_1 t_1)} p_{ii}^{(b_2 t_2)} \cdots p_{ii}^{(b_m t_m)} > 0.$$

Since the chain is irreducible, the claim follows by choosing t_0 sufficiently larger than T to allow all states to communicate with i. □

Let then $A \geq 0$ be an arbitrary nonnegative $n \times n$-matrix. Consider the set

$$\Lambda = \{\lambda \in \mathbb{R} \mid Ax \geq \lambda x \text{ for some } x \geq 0\}.$$

Clearly $0 \in \Lambda$, so $\Lambda \neq \varnothing$. Also, it is easy to see that the values in Λ are upper bounded by the maximal rowsum M of A. Thus $\Lambda \subseteq [0, M]$, and we may define

$$\lambda^* = \sup \Lambda.$$

To see that the supremum of Λ is actually attained by some $\lambda^* \in \Lambda$ and vector $x^* \geq 0$, observe that one may also define λ^* as

$$\lambda^* = \max_{x \in [0,1]^n} \min_{i=1,\dots,n} \frac{(Ax)_i}{x_i},$$

where in the case of $x_i = 0$, the quotient $\frac{(Ax)_i}{x_i}$ is defined as the appropriate limit to maintain continuity.

Theorem 1.7 (Perron-Frobenius) *For any strictly positive matrix $A > 0$ there exist a positive real number $\lambda^* > 0$ and a strictly positive vector $x^* > 0$ such that:*

(i) $Ax^* = \lambda^* x^*$;

(ii) *if $\lambda \neq \lambda^*$ is any other (in general complex) eigenvalue of A, then $|\lambda| < \lambda^*$;*

(iii) *λ^* has geometric and algebraic multiplicity 1.*

Proof. Define λ^* as above, and let $x^* \geq 0$ be a vector such that $Ax^* \geq \lambda^* x^*$. Since $A > 0$, also $\lambda^* > 0$.

(i) Suppose that it is not the case that $Ax^* = \lambda^* x^*$, i.e. that $Ax^* \geq \lambda^* x^*$, but not $Ax^* = \lambda^* x^*$. Consider the vector $y^* = Ax^*$. Since $A > 0$, $Ax > 0$ for any $x \gtrsim 0$; in particular now $A(y^* - \lambda^* x^*) = Ay^* - \lambda^* Ax^* = Ay^* - \lambda^* y^* > 0$, i.e. $Ay^* > \lambda^* y^*$; but this contradicts the definition of λ^*.

Consequently $Ax^* = \lambda^* x^*$, and furthermore $x^* = \frac{1}{\lambda^*} Ax^* > 0$.

(ii) Let $\lambda \neq \lambda^*$ be an eigenvalue of A and $y \neq 0$ the corresponding eigenvector, $Ay = \lambda y$. Denote $|y| = (|y_1|, \dots, |y_n|)$. Since $A > 0$, it is the case that

$$A|y| \geq |Ay| = |\lambda y| = |\lambda||y|.$$

By the definition of λ^*, it follows that $|\lambda| \leq \lambda^*$.

To prove strict inequality, let $\delta > 0$ be so small that the matrix $A_\delta = A - \delta I$ is still strictly positive. Then for any eigenvalue λ of A, $\lambda - \delta$ is an eigenvalue of A_δ and vice versa. Since $A_\delta > 0$, its largest eigenvalue is $\lambda^* - \delta$, i.e. for any other eigenvalue λ of A, $|\lambda - \delta| \leq \lambda^* - \delta$.

But this implies that A cannot have any eigenvalues $\lambda \neq \lambda^*$ on the circle $|\lambda| = \lambda^*$, because such would have $|\lambda - \delta| > |\lambda^* - \delta|$. (See Figure 5.)

(iii) We shall consider only the geometric multiplicity. Suppose there was another (real) eigenvector $y > 0$, linearly independent of x^*, associated to λ^*. Then one could form a linear combination $w = x^* + \alpha y$ such that $w \gtrsim 0$, but not $w > 0$. However, since $A > 0$, it must be the case that also $w = \frac{1}{\lambda^*} Aw > 0$. □

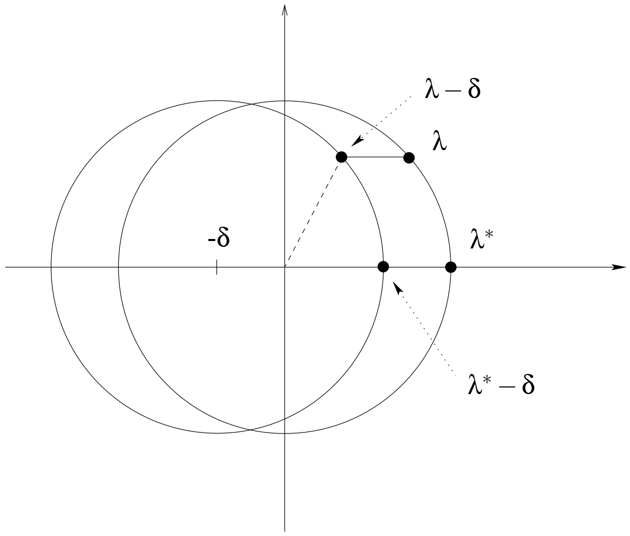

Figure 5: Maximality of the Perron-Frobenius eigenvalue.

Corollary 1.8 *If A is a nonnegative matrix ($A \geq 0$) such that some power of A is strictly positive ($A^n > 0$), then the conclusions of Theorem 1.7 hold also for A.* □

Note: In fact every nonnegative matrix $A \geq 0$ has a real "Perron-Frobenius" eigenvalue $\lambda^* \geq 0$ of maximum modulus, i.e. such that $|\lambda| \leq \lambda^*$ holds for all eigenvalues λ of A. But in this general case there may also be complex eigenvalues of equal modulus, and λ^* itself may be nonsimple, i.e. have multiplicity greater than one.

Proposition 1.9 *Let $A \geq 0$ be a nonnegative $n \times n$ matrix with row and column sums*

$$r_i = \sum_j a_{ij}, \quad c_j = \sum_i a_{ij}, \quad i, j = 1, \ldots, n$$

Then for the Perron-Frobenius eigenvalue λ^* of A the following bounds hold:

$$\min_i r_i \leq \lambda^* \leq \max_i r_i, \quad \min_j c_j \leq \lambda^* \leq \max_j c_j.$$

Proof. Let $x^* = (x_1, x_2, \ldots, x_n)$ be an eigenvector corresponding to λ^*, normalised so that $\sum_i x_i = 1$. Summing up the equations for $Ax^* = \lambda^* x^*$ yields:

$$
\begin{array}{ccccccccl}
a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = \lambda^* x_1 & \\
a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = \lambda^* x_2 & \\
 & & & & \vdots & & & & \\
a_{n1}x_1 & + & a_{n2}x_2 & + & \ldots & + & a_{nn}x_n & = \lambda^* x_n & \\
\hline
c_1x_1 & + & c_2x_2 & + & \ldots & + & c_nx_n & = \lambda^* \underbrace{(x_1 + \cdots + x_n)}_{1} & = \lambda^*
\end{array}
$$

Thus λ^* is a "weighted average" of the column sums, so in particular $\min_j c_j \leq \lambda^* \leq \max_j c_j$.

Applying the same argument to A^T, which has the same λ^* as A, yields the row sum bounds. □

Corollary 1.10 *Let $P \geq 0$ be the transition matrix of a regular Markov chain. Then there exists a unique distribution vector π such that $\pi P = \pi$ ($\Leftrightarrow P^T \pi^T = \pi^T$).*

Proof. By Lemma 1.6 and Corollary 1.8, P has a unique largest eigenvalue $\lambda^* \in \mathbb{R}$. By Proposition 1.9, $\lambda^* = 1$, because as a stochastic matrix all row sums of P (i.e. the column sums of P^T) are 1. Since the geometric multiplicity of λ^* is 1, there is a unique stochastic vector π (i.e. satisfying $\sum_i \pi_i = 1$) such that $\pi P = \pi$. □

1.3 Convergence of Regular Markov Chains

In Corollary 1.10 we established that a regular Markov chain with transition matrix P has a unique stationary distribution vector π such that $\pi P = \pi$.

By elementary arguments (page 3) we know that starting from any initial distribution q, if the iteration $q, qP, qP^2, \ldots$ converges, then it must converge to this unique stationary distribution.

However, it remains to be shown that if the Markov chain determined by P is regular, then the iteration always converges.

The following matrix decomposition is well known:

Lemma 1.11 (Jordan canonical form) *Let $A \in \mathbb{C}^{n \times n}$ be any matrix with eigenvalues $\lambda_1, \ldots, \lambda_l \in \mathbb{C}$, $l \leq n$. Then there exists an invertible matrix $U \in \mathbb{C}^{n \times n}$ such* *that*

$$UAU^{-1} = \begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_r \end{bmatrix}$$

where each J_i is a $k_i \times k_i$ ***Jordan block*** *associated to some eigenvalue λ of A:*

$$J_i = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}$$

The total number of blocks associated to a given eigenvalue λ corresponds to λ's geometric multiplicity, and their total dimension $\sum_i k_i$ to λ's algebraic multiplicity.
□

Now let us consider the Jordan canonical form of a transition matrix P for a regular Markov chain. Assume for simplicity that all the eigenvalues of P are real and distinct. (The general argument is similar, but needs more complicated notation.) Then the rows of U may be taken to be left eigenvectors of the matrix P, and the Jordan canonical form reduces to the familiar eigenvalue decomposition:

$$UPU^{-1} = \Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}.$$

In this case one notes that in fact the columns of $U^{-1} = V$ are precisely the *right* eigenvectors corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_n$. By Lemma 1.6 and Corollary 1.8, P has a unique largest eigenvalue $\lambda_1 = 1$, and the other eigenvalues may be ordered so that $1 > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$. The unique (up to normalisation) left eigenvector associated to eigenvalue 1 is the stationary distribution π, and the corresponding unique (up to normalisation) right eigenvector is $\mathbf{1} = (1, 1, \ldots, 1)$. If the first row of U is normalised to π, then the first column of V must be normalised to $\mathbf{1}$ because $UV = UU^{-1} = I$, and hence $(UV)_{11} = u_1 v_1 = \pi v_1 = 1$.

Denoting

$$\Lambda = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix},$$

we have then:

$$P^2 = (V\Lambda U)^2 = V\Lambda^2 U = V \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n^2 \end{bmatrix} U,$$

and in general

$$\begin{aligned} P^t &= V\Lambda^t U = V \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \lambda_2^t & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n^t \end{bmatrix} U \\ &\xrightarrow[t\to\infty]{} V \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 \end{bmatrix} U = \begin{bmatrix} v_{11}u_1 \\ v_{12}u_1 \\ \vdots \\ v_{1n}u_1 \end{bmatrix} = \begin{bmatrix} \pi \\ \pi \\ \vdots \\ \pi \end{bmatrix}. \end{aligned}$$

To make the situation even more transparent, represent a given initial distribution $q = q^0$ in the (left) eigenvector basis as

$$\begin{aligned} q &= \tilde{q}_1 u_1 + \tilde{q}_2 u_2 + \cdots + \tilde{q}_n u_n \\ &= \pi + \tilde{q}_2 u_2 + \cdots + \tilde{q}_n u_n, \qquad \text{where } \tilde{q}_i = \langle q^T, v_i \rangle = q v_i. \end{aligned}$$

Then

$$qP = (\pi + \tilde{q}_2 u_2 + \cdots + \tilde{q}_n u_n) P = \pi + \tilde{q}_2 \lambda_2 u_2 + \cdots + \tilde{q}_n \lambda_n u_n,$$

and generally

$$q^{(t)} = qP^t = \pi + \sum_{i=2}^{n} \tilde{q}_i \lambda_i^t u_i,$$

implying that $q^{(t)} \xrightarrow[t\to\infty]{} \pi$, and if the eigenvalues are ordered as assumed, then

$$||q^{(t)} - \pi|| = O(|\lambda_2|^t).$$

1.4 Transient Behaviour of General Chains

So what happens to the transient states in a reducible Markov chain?

A moment's thought shows that the transition matrix of an arbitrary (finite) Markov chain can be put in the following *canonical form*:

$$P = \left[\begin{array}{ccc|c} P_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & P_r & \\ \hline & & & \\ & R & & Q \\ & & & \end{array}\right]$$

where the r square matrices $P_1, \dots, P_r$ in the upper left corner represent the transitions within the r minimal closed classes, Q represents the transitions among transient states, and R represents the transitions from transient states to one of the closed classes.

In this ordering, stationary distributions (left eigenvectors of P corresponding to eigenvalue 1) must apparently be of the form $\pi = [\pi_1 \ \cdots \ \pi_r \quad 0 \ \cdots \ 0]$. (Note that since Q has at least one row sum less than 1, by the proof argument in Proposition 1.9 also all of its eigenvalues have modulus less than 1. Thus the only solution of the stationarity equation $\mu Q = \mu$ is $\mu = 0$.)

Consider then the *fundamental matrix* $M = (I-Q)^{-1}$ of the chain. Intuitively, if M is well-defined, it corresponds to $M = I + Q + Q^2 + \dots$, and represents all the possible transition sequences the chain can have without exiting Q.

Theorem 1.12 *For any finite Markov chain with transition matrix as above, the matrix $I - Q$ is invertible, and its inverse can be represented as the convergent series $M = I + Q + Q^2 + \dots$*

Proof. Since for any $t \geq 1$,

$$(I-Q)(I+Q+\cdots+Q^{t-1}) = I - Q^t,$$

and $Q^t \to 0$ as $t \to \infty$, the result follows. □

A transparent stochastic interpretation of the fundamental matrix may be obtained by considering any two transient states i, j in a Markov chain as above. Then:

$$\Pr(X_t = j \mid X_0 = i) = Q^t_{ij} \triangleq q^{(t)}_{ij}.$$

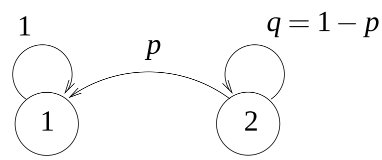

Figure 6: A Markov chain representing the geometric distribution.

Thus,

$$\begin{aligned} E[\text{number of visits to } j \in T \mid X_0 = i \in T] &= q_{ij}^{(0)} + q_{ij}^{(1)} + q_{ij}^{(2)} + \dots \\ &= I_{ij} + Q_{ij} + Q_{ij}^2 + \dots \\ &= M_{ij} \triangleq m_{ij}. \end{aligned}$$

Furthermore,

$$\begin{aligned} & E[\text{number of moves in } T \text{ before exiting to } C \mid X_0 = i \in T] \\ = & \sum_{j \in T} E[\text{number of visits to } j \in T \mid X_0 = i \in T] \\ = & \sum_{j \in T} m_{ij} \\ = & (M\mathbf{1})_i. \end{aligned}$$

As another application, let b_{ij} be the probability that the chain when started in transient state $i \in T$ will enter a minimal closed class via state $j \in C$. Denote $B = (b_{ij})_{i \in T, j \in C}$. Then in fact $B = MR$.

Proof. For given $i \in T, j \in C$,

$$b_{ij} = p_{ij} + \sum_{k \in T} p_{ik} b_{kj}.$$

Thus,

$$B = R + QB \quad \Rightarrow \quad B = (I - Q)^{-1} R = MR.$$

Example 1.4 *The geometric distribution.* Consider the chain of Figure 6, arising e.g. from biased coin-flipping The transition matrix in this case is

$$P = \begin{bmatrix} 1 & 0 \\ p & q \end{bmatrix}.$$

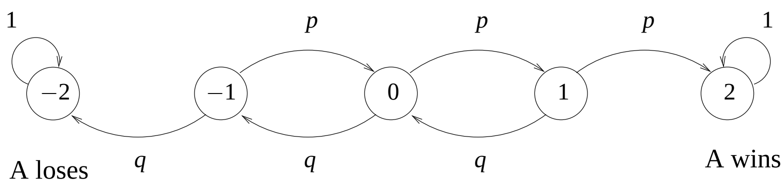

Figure 7: A Markov chain representing a coin-flipping game.

Now $Q = (q)$, $M = (1-q)^{-1} = 1/p$. Thus, e.g.

$$E[\text{number of visits to 2 before exiting to 1} \mid X_0 = 2] = M\mathbf{1} = \frac{1}{p}.$$

An elementary way to obtain the same result would be:

$$\begin{aligned} E[\text{number of visits}] &= \sum_{k\geq 0} \Pr[\text{number of visits} = k] \cdot k \\ &= \sum_{k\geq 0} \Pr[\text{number of visits} \geq k] \\ &= 1 + q + q^2 + \cdots = \frac{1}{1-q} = \frac{1}{p}. \end{aligned}$$

Example 1.5 *Gambling tournament.* Players A and B toss a biased coin with A's success probability equal to p and B's success probability equal to $1-p=q$. The person to first obtain n successes over the other wins. What are A's chances of winning, given that he initially has k successes over B, $-n \leq k \leq n$? (A more technical term for this process is "one-dimensional random walk with two absorbing barriers.")

For simplicity, let us consider only the case $n = 2$. Then the chain is as represented in Figure 7, with transition matrix:

	−2	−1	0	1	2
−2	1	0	0	0	0
−1	q	0	p	0	0
0	0	q	0	p	0
1	0	0	q	0	p
2	0	0	0	0	1

i.e. in canonical form:

	−2	2	−1	0	1
−2	1	0	0	0	0
2	0	1	0	0	0
−1	q	0	0	p	0
0	0	0	q	0	p
1	0	p	0	q	0

Thus, $M = (I-Q)^{-1}$

$$= \begin{bmatrix} 1 & -p & 0 \\ -q & 1 & -p \\ 0 & -q & 1 \end{bmatrix}^{-1} = \frac{1}{p^2+q^2} \begin{bmatrix} p+q^2 & p & p^2 \\ q & 1 & p \\ q^2 & q & q+p^2 \end{bmatrix}$$

and so $B = MR$

$$= \frac{1}{p^2+q^2} \begin{bmatrix} p+q^2 & p & p^2 \\ q & 1 & p \\ q^2 & q & q+p^2 \end{bmatrix} \begin{bmatrix} q & 0 \\ 0 & 0 \\ 0 & p \end{bmatrix} = \frac{1}{p^2+q^2} \begin{bmatrix} qp+q^3 & p^3 \\ q^2 & p^2 \\ \underbrace{q^3}_{\text{A loses}} & \underbrace{pq+p^3}_{\text{A wins}} \end{bmatrix}.$$

1.5 Reversible Markov Chains

We now introduce an important special class of Markov chains often encountered in algorithmic applications. Many examples of these types of chains will be encountered later.

Intuitively, a "reversible" chain has no preferred time direction at equilibrium, i.e. any given sequence of states is equally likely to occur in forward as in backward order.

A Markov chain determined by the transition matrix $P = (p_{ij})_{i,j \in S}$ is *reversible* if there is a distribution π that satisfies the *detailed balance* conditions:

$$\pi_i p_{ij} = \pi_j p_{ji} \quad \forall\, i, j \in S.$$

Theorem 1.13 *A distribution satisfying the detailed balance conditions is stationary.*

Proof. It suffices to show that, assuming the detailed balance conditions, the following stationarity condition holds for all $i \in S$:

$$\pi_i = \sum_{j \in S} \pi_j p_{ji}.$$

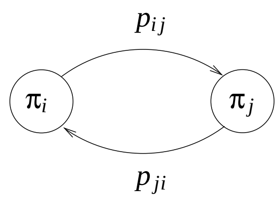

Figure 8: Detailed balance condition $\pi_i p_{ij} = \pi_j p_{ji}$.

But this is straightforward:

$$\sum_{j\in S} \pi_j p_{ji} = \sum_{j\in S} \pi_i p_{ij} = \pi_i \sum_{j\in S} p_{ji} = \pi_i.$$

□

Observe the intuition underlying the detailed balance condition: At stationarity, an equal amount of probability mass flows in each step from i to j as from j to i.(The "ergodic flows"' between states are in pairwise balance; cf. Figure 8.)

Example 1.6 *Random walks on graphs.*

Let $G = (V,E)$ be a (finite) graph, $V = \{1,\dots,n\}$. Define a Markov chain on the nodes of G so that at each step, one of the current node's neigbours is selected as the next state, uniformly at random. That is,

$$p_{ij} = \begin{cases} \frac{1}{d_i}, & \text{if } (i,j) \in E \\ 0, & \text{otherwise} \end{cases} \qquad (d_i = \deg(i))$$

Let us check that this chain is reversible, with stationary distribution

$$\pi = \begin{bmatrix} \frac{d_1}{d} & \frac{d_2}{d} & \cdots & \frac{d_n}{d} \end{bmatrix},$$

where $d = \sum_{i=1}^{n} d_i = 2|E|$. The detailed balance condition is easy to verify:

$$\pi_i p_{ij} = \begin{cases} \frac{d_i}{d} \cdot \frac{1}{d_i} = \frac{1}{d} = \frac{d_j}{d} \cdot \frac{1}{d_j} = \pi_j p_{ji}, & \text{if } (i,j) \in E \\ 0 = \pi_j p_{ji}, & \text{if } (i,j) \notin E \end{cases}$$

Example 1.7 *A nonreversible chain.*

Consider the three-state Markov chain shown in Figure 9. It is easy to verify that this chain has the unique stationary distribution $\pi = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$. However, for any $i = 1,2,3$:

$$\pi_i p_{i,(i+1)} = \frac{1}{3} \cdot \frac{2}{3} = \frac{2}{9} > \pi_{i+1} p_{(i+1),i} = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}.$$

Thus, even in a stationary situation, the chain has a "preference" of moving in the counter-clockwise direction, i.e. it is not time-symmetric.

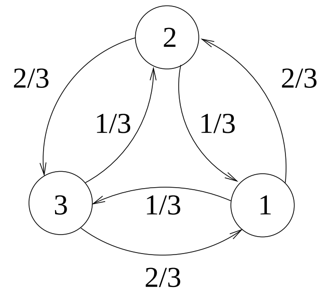

Figure 9: A nonreversible Markov chain.

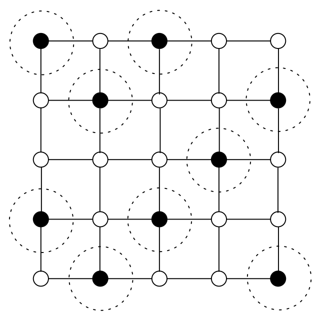

Figure 10: Hard-core colouring of a lattice.

2 Markov Chain Monte Carlo Sampling

We now introduce *Markov chain Monte Carlo (MCMC) sampling*, which is an extremely important method for dealing with "hard-to-access" distributions.

Assume that one needs to generate samples according to a probability distribution π, but π is too complicated to describe explicitly. A clever solution is then to construct a Markov chain that converges to stationary distribution π, let it run for a while and then sample states of the chain. (However, one obvious problem that this approach raises is determining how long is "for a while"? This leads to interesting considerations of the convergence rates and "rapid mixing" of Markov chains.)

Example 2.1 *The hard-core model.*

A *hard-core colouring* of a graph $G = (V, E)$ is a mapping

$$\xi : V \to \{0, 1\} \qquad \text{("empty" vs. "occupied" sites)}$$

such that

$$(i, j) \in E \;\Rightarrow\; \xi(i) = 0 \vee \xi(j) = 0 \qquad \text{(no two occupied sites are adjacent)}$$