DEMONSTRATIO MATHEMATICA
Vol. XXX No 4 1997

Günther Karigl

ON A GRAPH THEORETIC DESCRIPTION OF REGULAR MARKOV CHAINS AND POSITIVE LINEAR SYSTEMS

Abstract. Regular Markov chains are widely used in stochastic modelling and exhibit a well known limiting behaviour. This is also true for a certain class of positive linear systems according to the Perron–Frobenius theorem. In our paper we give a characterization of regularity by graph theoretic properties. Moreover, we describe all regular Markov chains with given transition and limiting behaviour.

1. Introduction

Markov chain models represent an important class of stochastic dynamic systems. Consider a finite Markov chain with state set $S = \{S_1, \ldots, S_n\}$, $n \in \mathrm{N}$, and stationary transition probabilities $p_{ij} \geq 0$, $i, j = 1, \ldots, n$. Whereas onestep transitions from state S_i to state S_j are determined by the transition matrix $P = (p_{ij})$, the probabilities for m-step transitions $p_{ij}^{(m)}$, $m \in \mathrm{N}$, are given by the elements of P^m. In order to obtain a classification of Markov chains a binary relation ϱ on the state set S can be defined by

$$S_i \varrho S_j \Leftrightarrow i = j \text{ or } p_{ij}^{(m)} > 0 \text{ for some positive integer } m,$$

i.e. $S_i \varrho S_j$ holds iff S_j can be reached from S_i. From the algebraic point of view ϱ is a quasi-ordering relation and implies an equivalence relation θ on S according to

$$S_i \theta S_j \Leftrightarrow (S_i \varrho S_j) \text{ and } (S_j \varrho S_i).$$

The corresponding partition divides the set of all states into communicating classes. A special case occurs if there is only one class whose states are all aperiodic, i.e. every state in S can be reached from every other state by a fixed number of transitions. In this case the Markov chain is called regular (see e.g. [3], [5] for more details).

Regular Markov chains have a wide variety of applications. Let us consider a discrete stochastic process $x_{t+1} = x_t P$ for $t = 0, 1, 2, \ldots$ where x_t

denotes the probability distribution vector of the process and P is the transition matrix corresponding to a regular Markov chain. Then according to the basic limit theorem for Markov chains, the long-term probabilities converge to a limiting distribution which is independent of the initial conditions. To be more precise, $\lim_{t\to\infty} x_t = a$, where a is a left eigenvector of P corresponding to the eigenvalue $\lambda = 1$. Moreover, the matrix P^t approaches a matrix each of whose rows is equal to a, if $t \to \infty$. Our aim is to give a description of regularity by graph theoretic properties.

In close connection to Markov chains there are positive linear systems, i.e. dynamic systems in which the state variables are always positive (or at least nonnegative). These arise frequently in many real systems. Let us consider a discrete dynamic system $x_{t+1} = Ax_t$, where x_t denotes the state vector at time t and A is real $n \times n$-matrix with $A \geq 0$. If A is strictly positive, i.e. $A > 0$, or more generally $A \geq 0$ and $A^m > 0$ for some $m \in \mathrm{N}$, according to the Perron—Frobenius theorem (cf. [5]) there exists a dominant eigenvalue λ af A of largest absolute value which is in fact positive and simple, and a corresponding positive eigenvector a. Furthermore the state vector x_t tends to be aligned with $\lambda^t a$ as $t \to \infty$. Thus the system approaches a stable relative distribution. In section 2 we give a description of positive systems which behave in that way.

Moreover, we characterize regular Markov chains with transition probabilities proportional to given transition numbers and with a given limiting distribution in section 3. This problem is of interest in the context of random walk models and cellular automata.

2. A characterization of graphs corresponding to regular Markov chains

Let $P = (p_{ij})$ be the transition matrix of a finite Markov chain with state set $S = \{S_1, \ldots, S_n\}$ and define a corresponding directed graph $G_P = (V, E)$ with vertex set $V = \{1, \ldots, n\}$ and edge set $E \subseteq V^2$ such that

$$ij \in E \Leftrightarrow p_{ij} > 0 \tag{1}$$

(where ij denotes the edge (i, j)). Then clearly state S_j can be reached from state S_i iff there exists a directed walk in G_P which connects vertex i to vertex j. The communicating classes of the Markov chain correspond to the connected components of G. In particular P is regular iff any two vertices in G_P can be connected by a directed walk of fixed length.

On the other hand, let $A \geq 0$ be any nonnegative real $n \times n$-matrix and define a directed graph $G_A = (V, E)$ with vertex set $V = \{1, \ldots, n\}$ and edge set $E \subseteq V^2$ such that

$$ij \in E \Leftrightarrow a_{ij} > 0 \tag{2}$$

just as above. Then the elements of the matrix A^m correspond to the directed walks of length m in G_A ($m \in \mathbb{N}$), and $A^m > 0$ for some positive integer m holds iff again two vertices in G_A can be connected by a directed walk of fixed length.

If the transition matrix P or the matrix A is symmetric, the corresponding graphs may be considered as nondirected instead of directed graphs. The question whether a Markov chain is regular or not, the question whether a positive linear system approaches a stable distribution or not, carries over to the problem under which conditions any two vertices in the corresponding (directed or nondirected) graph can be connected by a walk of some fixed length. This question is answered by the following two theorems.

THEOREM 2.1. *Let $G = (V, E)$ be a finite, nondirected graph. Then the following conditions are equivalent:*

(i) *There exists a positive integer $m \in \mathbb{N}$ such that any two vertices $v, w \in V$ can be connected by a walk of fixed length m.*

(ii) *G is connected and has a cycle of odd length.*

Observe that if condition (i) holds for any particular m, then it also holds for all m' with $m' \geq m$. Furthermore, condition (ii) holds if G is connected and there exists at least one edge $vv \in E$ for some $v \in V$.

Proof: (i) ⇒ (ii) Condition (i) clearly implies that G is connected. Next let us assume that all proper cycles and hence all loops have even length. Then for any $vw \in E$ let p_1 and p_2 be two walks of length m connecting v to v and w to v, respectively. Then $p = p_1 + vw + p_2$ forms a loop of length $2m + 1$ which is a contradiction.

(ii) ⇒ (i) Let p_1 be any cycle of odd length $l(p_1) = l_1$ passing through $u_1 \in V$ and let $p_2 = u_1u_2u_1$ be a loop of length 2. Then by combining p_1 and p_2, loops of arbitrary length $\geq l_1$ through u_1 can be generated.

Now let $v, w \in V$ be any two vertices and let d_G be diameter of graph G. Since G is connected there exist walks vu_1 and u_1w connecting v to u_1 and u_1 to w, respectively, of lengths $l(vu_1) \leq d_G$, $l(u_1v) \leq d_G$. Moreover, there exists a loop u_1u_1 of length

$$l(u_1u_1) = l_1 + (d_G - l(vu_1)) + (d_G - l(u_1w)) \geq l_1$$

as stated before. It follows that $p = vu_1 + u_1u_1 + u_1w$ represents a walk which connects v to w and has length $l(p) = 2d_G + l_1 =: m$, where m does not depend on the vertices v and w, as claimed in (i). ∎

THEOREM 2.2. *Let $G = (V, E)$ be a finite, directed graph. Then the following conditions are equivalent:*

(i) *There exists a positive integer $m \in \mathbb{N}$ such that any two vertices $v, w \in V$ can be connected by a directed walk of fixed length m.*

(ii) *G is strongly connected and there exist two directed loops p_1, p_2 such that $gcd(l(p_1), l(p_2)) = 1$.*

All remarks made subsequent to Theorem 2.1 also hold in the case of directed graphs. In addition we note that condition (ii) clearly holds if there are two proper cycles p_1, p_2 in G such that $gcd(l(p_1), l(p_2)) = 1$; the latter condition, however, is not necessary (e.g. consider a graph G which consists of three connected cycles of lengths 6, 10 and 15, respectively). It seems to be reasonable to replace condition (ii) by the demand that G is strongly connected and the greatest common divisor of the lengths of all cycles in G equals to 1. This, however, remains as an open question.

Proof: (i) ⇒ (ii) Obviously G is strongly connected if condition (i) holds. In order to prove the existence of two directed loops with relative prime lengths take any edge $vw \in E$ and choose two walks $p_1 = vv$ from v to v and $p_2' = wv$ connecting w to v according to (i) such that $l(p_1) = l(p_2') = m$. Then p_1 and $p_2 = vw + p_2'$ both are loops and

$$gcd(l(p_1), l(p_2)) = gcd(m, m+1) = 1.$$

(ii) ⇒ (i) Let p_1 and p_2 be two directed loops passing through $u_1, u_2 \in V$, respectively, of lengths $l(p_1) = l_1$, $l(p_2) = l_2$ with $gcd(l_1, l_2) = 1$. Then there exist $x_0, y_0 \in \mathbf{Z}$ such that $l_1 x_0 + l_2 y_0 = 1$. We consider the Diophantine equation

$$l_1 x + l_2 y = m' \quad \text{for } m' \in \mathbf{N} \tag{3}$$

and ask for solution $x, y \geq 0$. The solutions of equation (3) in $\mathbf{Z}$ are given by

$$x = m'x_0 + kl_2, \; y = m'y_0 - kl_1, \; k \in \mathbf{Z}$$

whence $x \geq 0$, $y \geq 0$ iff k belongs to the interval $-m'x_0/l_2 \leq k \leq m'y_0/l_1$ of length $m'/(l_1 l_2)$. So if $m' \geq l_1 l_2$ this interval contains at least one integer k and hence nonnegative solutions of equation (3) exist. (In fact equation (3) has nonnegative solutions if $m' \geq (l_1 - 1)(l_2 - 1)$. This is known as the postage stamps problem [6] which is also connected to the notion of Frobenius numbers of certain semigroups [4]). Let x and y be solutions of equation (3) with

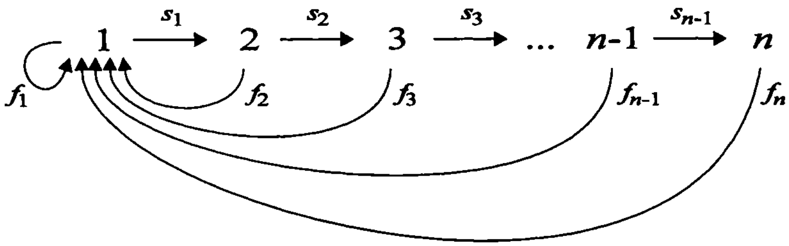
$$m' \geq l_1 l_2 + (d_G - l(u_1 u_2)) + (d_G - l(u_2 u_1)) \geq l_1 l_2.$$

It follows that $p = xp_1 + u_1 u_2 + yp_2 + u_2 u_1$ forms a directed loop of arbitrary length $l(p) \geq l_1 l_2 + 2d_G$ passing through u_1 (where xp_1 denotes the loop generated by x repetitions of p_1, and yp_2 is defined in the same way). Therefrom we conclude just as in the nondirected case that any two vertices $v, w \in V$ can be connected by a directed walk of length $m = l_1 l_2 + 4d_G$. This completes the proof of Theorem 2.2. ∎

In order to give an example let us consider the cohort population model of population biology (cf. [5] and [7]) describing the growth of a population which is divided into age groups according to

$$(4)\qquad x_{t+1} = Ax_t \text{ with } A = \begin{pmatrix} f_1 & f_2 & \cdots & f_n \\ s_1 & 0 & \ldots 0 & \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ldots & s_{n-1} & 0 \end{pmatrix}, \quad t = 0,1,2,\ldots$$

where x_t denotes the population distribution at time t, $f_i \geq 0$ is the fertility rate and $s_i > 0$ is the survival rate of the i-th age group for $i = 1,\ldots,n$. The graph G_A corresponding to A looks as follows:



According to Theorem 2.2 the population approaches a stable age distribution, determined by the dominant eigenvector of A, if and only if there are two age groups i and j such that $f_i > 0$, $f_j > 0$ and $gcd(i,j) = 1$. This is in particular true if $f_i > 0$ and $f_{i+1} > 0$ for some $i < n$.

3. Transitions in Markov chain models

In this last section we consider an important question concerning the construction of Markov chain models. In many applications, e.g. in the context of random walks or cellular automata (cf. [1], [2]), transition probabilities are expected to be proportional to the lengths (or areas) of common boundaries between cells, and limiting probabilities are assumed to be proportional to the areas (or volumes) of cells. In general, we are interested in Markov chains with given transition and limiting behaviour. In the sequel a characterization of these processes is given.

First of all we introduce the notion of a proportional matrix belonging to a Markov chain. Let $K = (k_{ij})$ be any real $n \times n$-matrix such that

$$(5)\qquad k_{ij} \geq 0 \text{ for } i \neq j \text{ and } k_{ii} = -\sum_{j\neq i} k_{ij}, \quad i,j = 1,\ldots,n.$$

Then K is called a proportional matrix of a Markov chain on $\{1,\ldots,n\}$ with transition matrix $P = (p_{ij})$, if

$$(6)\qquad \forall i = 1,\ldots,n \; \exists c_i > 0 : \; p_{ij} = c_i k_{ij} \;\text{ for all } j \neq i,$$

i.e. each probability p_{ij} of transition from state i to state j is proportional to k_{ij}.

THEOREM 3.1. *All Markov chains with proportional matrix K are given by*

$$P_K(c_1,\ldots,c_n) = E + \mathrm{diag}(c_1,\ldots,c_n)K, \tag{7}$$

where $0 < c_i \le (\sum_{j\ne i} k_{ij})^{-1}$ *(if* $\sum_{j\ne i} k_{ij} > 0$*) for all i and* $\mathrm{diag}(c_1,\ldots,c_n)$ *is the $n \times n$-diagonal matrix with diagonal elements $c_1,\ldots,c_n$.*

The proof of this result is obvious from (5), (6) and the fact that P_K is a stochastic matrix (i.e. $p_{ij} \ge 0$ and $\sum_j p_{ij} = 1$ for all i,j). ∎

Now let $K = (k_{ij})$ be a matrix as defined in (5) and let $a = (a_i) > 0$ be a positive n-dimensional vector. Furthermore, let $G_K = (V,E)$ be the directed graph with vertex set $V = \{1,\ldots,n\}$ and edges $E \subseteq V^2$ such that

$$ij \in E \Leftrightarrow k_{ij} > 0.$$

The following result gives a description of all regular Markov chains with transition behaviour determined by K and a limiting distribution proportional to vector a.

THEOREM 3.2. *Let K be any matrix as defined in (5) and let $a > 0$ be any positive vector. Then there exists a regular Markov chain with proportional matrix K iff the graph G_K corresponding to K is strongly connected. In this case K has exactly one left eigenvector $u = (u_i) > 0$ corresponding to the eigenvalue $\lambda = 0$, and all regular Markov chains with proportional matrix K and limiting distribution proportional to vector a are given by*

$$P_{K,a}(c) = E + c\,\mathrm{diag}\left(\frac{u_1}{a_1},\ldots,\frac{u_n}{a_n}\right)K, \tag{8}$$

where $0 < c \le \min_i \frac{a_i}{u_i}(\sum_{j\ne i} k_{ij})^{-1}$.

Proof. If G_K is not strongly connected this is also true for any graph G_P corresponding to any Markov chain P with proportional matrix K, and P cannot be regular by Theorem 2.2.

If, however, G_K has only one communicating class then there exists a nonsingular matrix C such that $P = E + CK$ is regular. Observe that for any left or right eigenvector of P belonging to the eigenvalue $\lambda = 1$ there exists a corresponding eigenvector of K with respect to the eigenvalue $\lambda = 0$ and vice versa, since

$$Px = x \Leftrightarrow Kx = 0 \quad\text{and}\quad yP = y \Leftrightarrow (yC)K = 0.$$

In particular, for any regular P there exists exactly one left eigenvector $u > 0$ of K corresponding to $\lambda = 0$.

Now let P be any regular Markov chain with proportional matrix K and limiting distribution proportional to a, i.e. $aP = a$. Then $P = P_K(c_1, \dots, c_n)$ by Theorem 3.1 and

$$aP = a(E + \operatorname{diag}(c_1, \dots, c_n)K) = a \Rightarrow a \operatorname{diag}(c_1, \dots, c_n)K = 0.$$

Therefore, $a \operatorname{diag}(c_1, \dots, c_n)$ is a left eigenvector of K corresponding to $\lambda = 0$ and hence proportional to u. It follows that $c_i = c(u_i/a_i)$ for $i = 1, \dots, n$ and some $c > 0$. Thus we have

$$P = P_{K,a}(c) = E + c \operatorname{diag}\left(\frac{u_1}{a_1}, \dots, \frac{u_n}{a_n}\right)K$$

where $0 < c \leq \min_i \frac{a_i}{u_i}(\sum_{j \neq i} k_{ij})^{-1}$ as claimed in (8). (The regularity of P guarantees that $\sum_{j \neq i} k_{ij} > 0$ for any i).

Conversely, any regular Markov chain $P_{K,a}(c)$ has a limiting distribution proportional to vector a since

$$aP_{K,a}(c) = a(E + c \operatorname{diag}\left(\frac{u_1}{a_1}, \dots, \frac{u_n}{a_n}\right)K) = a + cuK = a$$

which completes the proof of Theorem 3.2. ■

Finally let us consider a regular Markov chain with symmetric proportional matrix K and a uniform limiting distribution $a = (1, \dots, 1)$ which implies $u = (1, \dots, 1)$. This is a reasonable assumption in the context of certain cellular automata models where transitions from one cell to another do not depend on the direction of movement, and eventually each cell can be reached with the same probability. According to (8) it follows that

$$P_{K,1}(c) = E + cK = (p_{ij}) \text{ with } p_{ij} = ck_{ij} \text{ for } i \neq j \text{ and } p_{ii} = 1 - c\sum_{j \neq i} k_{ij}$$

with $0 < c \leq (\max_i \sum_{j \neq i} k_{ij})^{-1}$. If we set $p = c \max_i \sum_{j \neq i} k_{ij}$ then $0 < p \leq 1$ and

$$p_{ij} = pk_{ij}\Big(\max_i \sum_{j \neq i} k_{ij}\Big)^{-1}$$

is the probability of a transition from state i to state j $(i \neq j)$ whereas

$$p_{ii} = 1 - p\sum_{j \neq i} k_{ij}\Big(\max_i \sum_{j \neq i} k_{ij}\Big)^{-1}$$

is the probability that no transition from state i to any other state occurs. Thus, $1 - p \geq 0$ can be interpreted as the minimal probability of remaining within any state.

References

[1] D. Dorninger, G. Karigl and J. Loidl, *Simulation of chromosomal homology searching in meiotic pairing*, J. Theor. Biol. 176 (1995) 247—260.

[2] D. Dorninger, G. Karigl and J. Loidl, *A cellular automaton model for chromosome interlocking in meiotic pairing*, Simul. Pract. Theor. (1997). To appear.

[3] W. Feller, *An Introduction to Probability Theory and its Applications, Volume 1.* John Wiley and Sons, New York, 1968.

[4] R. Fröberg, *The Frobenius number of some semigroups*, Comm. Algebra 22 (1994) 6021–6024.

[5] D. G. Luenberger, *Introduction to Dynamic Systems. Theory, Models and Applications*, John Wiley and Sons, New York, 1979.

[6] A. R. Meijer, *Postage stamps and the generators of a cyclic group*, Internat. J. Math. Ed. Sci. Tech. 22 (1991). 683-693.

[7] W. Nöbauer und W. Timischl, *Mathematische Modelle in der Biologie*, Friedrich Vieweg&Sohn, Braunschweig, 1979.

DEPARTMENT OF ALGEBRA AND DISCRETE MATHEMATICS
VIENNA UNIVERSITY OF TECHNOLOGY
Wiedner Hauptstr. 8–10/118
A-1040 VIENNA, AUSTRIA
E-mail: G.Karigl@tuwien.ac.at

Received June 16, 1997.