2.6 Markov Chains

Markov chains are stochastic processes in which the probability that the process will be in state *j* at time *t* depends only on the state *i* at time *t-1*. A "state" of a finite-population GA is simply a particular finite population. The behavior of certain GAs may be modeled by a Markov chain. In what follows we give a brief description of the basic concepts.

2.6.1 Definitions and Theorems[1]

FIXED VECTORS

If $\vec{u}$ is a vector with *n* components we call $\vec{u} \neq 0$ a fixed vector (or *fixed point*) of **A** if $\vec{u}$ is left "fixed" (not changed) when multiplied by **A**: $\vec{u}\mathbf{A} = \vec{u}$.

THEOREM 2.6.1.1

If $\vec{u}$ is a fixed vector of matrix **A**, then every nonzero scalar multiple $k\vec{u}$ is also a fixed vector of **A**.

EXAMPLE E.2.6.1.1

Let $A = \begin{pmatrix} 2 & 1 \\ 2 & 3 \end{pmatrix}$. Then the vector $\vec{u} = (2, -1)$ is a fixed point of **A**. For,

$$\vec{u}\mathbf{A} = (2,-1)\begin{pmatrix} 2 & 1 \\ 2 & 3 \end{pmatrix} = (2\cdot 2 - 1\cdot 2, 2\cdot 1 - 1\cdot 3) = (2, -1) = \vec{u}.$$

PROBABILITY VECTOR

A vector $\vec{u} = (u_1, \ldots, u_n)$ is called a *probability vector* if its components are nonnegative and their sum is 1.

STOCHASTIC MATRICES

A square matrix $\mathbf{P} = (p_{ij})$ is called a *stochastic matrix* if each of its rows is a probability vector.

THEOREM 2.6.1.2

If **A** and **B** are stochastic matrices then the product **AB** is a stochastic matrix.

---

[1] The reader may consult [FELL70].

COROLLARY 2.6.1.2

All powers $\mathbf{A}^n$ of a stochastic matrix $\mathbf{A}$ are stochastic matrices.

REGULAR MATRICES

A stochastic matrix $\mathbf{P}$ is said to be *regular* if all the entries of some power $\mathbf{P}^m$ are positive.

EXAMPLE E.2.6.1.3

The stochastic matrix $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ ½ & ½ \end{pmatrix}$ is regular since $\mathbf{A}^2 = \begin{pmatrix} ½ & ½ \\ ¼ & ¾ \end{pmatrix}$ is positive in every entry.

THEOREM 2.6.1.3

Let $\mathbf{P}$ be a regular stochastic matrix. Then:

a) $\mathbf{P}$ has a unique fixed probability vector $\vec{t}$, and the components of $\vec{t}$ are all positive.
b) The sequence $\mathbf{P}$, $\mathbf{P}^2$, ... of powers of $\mathbf{P}$ approaches the matrix $\mathbf{T}$ whose rows are all equal to the fixed point $\vec{t}$.
c) If $\vec{p}$ is any probability vector, then the sequence of vectors $\vec{p}\mathbf{P}$, $\vec{p}\mathbf{P}^2$, ... approaches the fixed point $\vec{t}$.

EXAMPLE E.2.6.1.4

Consider the regular stochastic matrix $\mathbf{P} = \begin{pmatrix} 0 & 1 \\ ½ & ½ \end{pmatrix}$. We seek a probability vector which we can denote by $\vec{t} = (x, 1 - x)$, such that $\vec{t}\mathbf{P} = \vec{t}$. Thus,

$$(x, 1-x)\begin{pmatrix} 0 & 1 \\ ½ & ½ \end{pmatrix} = (x, 1-x)$$

from which we obtain

$$(½ - ½x,\ ½ + ½x) = (x,\ 1 - x) \text{ or } \begin{cases} ½ - ½x = x \\ ½ + ½x = 1 - x \end{cases} \text{ or } x = \frac{1}{3} \rightarrow \vec{t} = \left(\frac{1}{3}, \frac{2}{3}\right) \approx (.33,\ .67)$$

This is the unique fixed point of $\mathbf{P}$. From theorem 2.6.1.3., the sequence $\mathbf{P}$, $\mathbf{P}^2$,... approaches $\mathbf{T}$, whose rows are each equal to the fixed point $\vec{t}$.

$$T = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} \approx \begin{pmatrix} .33 & .67 \\ .33 & .67 \end{pmatrix} \rightarrow \vec{t} \approx (.33,\ .67)$$

DEFINITION

A square matrix **A** : n $\times$ n is said to be:

a) *Positive* (**A** > **0**), if $a_{ij} > 0$ for $i, j \in \{1, \ldots, n\}$.

b) *Nonnegative* (**A** ≥ **0**), if $a_{ij} \geq 0$ for all $i, j \in \{1, \ldots, n\}$.

A nonnegative matrix is said to be:

c) *Regular*, if there exists a $k \in \mathbb{N}$ such that $\mathbf{A}^k$ is positive.

d) *Reducible*, if **A** can be brought into the form

$$\begin{pmatrix} \boldsymbol{C} & \mathbf{0} \\ \boldsymbol{R} & \boldsymbol{T} \end{pmatrix}$$

by applying the same permutations to rows and columns[2].

e) *Irreducible*, if it is nor reducible.

f) *Stochastic*, if $\sum_{j=1}^{n} a_{ij} = 1$ for all $i \in \{1, \ldots, n\}$.

A stochastic matrix is said to be:

g) *Stable*, if it has identical rows.

h) *Column allowable* (*c-allowable*), if it has at least one positive entry in each column.

□

Note that every positive matrix is also regular.

LEMMA 2.6.1

Let **C**, **M** and **S** be stochastic matrices, where **M** is positive and **S** is c-allowable. Then the product **CMS** is positive.

THEOREM 2.6.1.4

Let **P** be a reducible stochastic matrix, where **C** : m $\times$ m is a regular stochastic matrix and **R**, **T** $\neq$ **0**. Then

$$P^{\infty} = \lim_{k \to \infty} P^{k} = \lim_{k \to \infty} \begin{pmatrix} C^{k} & 0 \\ \sum_{i=0}^{k-1} T^{i} R C^{k-i} & T^{k} \end{pmatrix} = \begin{pmatrix} C^{\infty} & 0 \\ R^{\infty} & 0 \end{pmatrix}$$

---

[2] **C** and **T** are square matrices.

is a stable stochastic matrix with $\mathbf{P}^{\infty} = \mathbf{I}' \vec{p}^{\infty}$, where $\vec{p}^{\infty} = \vec{p}^{0}\,\mathbf{P}^{\infty}$ is unique regardless of the initial configuration and $\vec{p}^{\infty}$ satisfies: $\vec{p}_i^{\,\infty} > 0$ for $1 \le i \le m$ and $\vec{p}_i^{\,\infty} = 0$ for $m < i \le n$.

2.6.2 Finite Markov Chains

Consider a sequence of trials whose outcomes $X_1, X_2, \ldots$ satisfy:

a) Each outcome belongs to a *finite* set of outcomes $\{a_1, a_2, \ldots, a_m\}$ called the *state space* of the system. If the outcome of the n-th trial is a_i we say that the system is in state a_i at time n or at the n-th step.
b) The outcome of any trial depends, at most, upon the outcome of the immediately preceding trial, i.e. there is a given probability p_{ij} that a_j occurs immediately after a_i occurs for every pair (a_i, a_j).

Such a stochastic process is called a finite Markov chain. The numbers p_{ij} are called the *transition probabilities* and may be arranged in a matrix called the *transition matrix*, thus:

$$\mathbf{P} = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ p_{m1} & p_{m2} & \cdots & p_{mn} \end{pmatrix}$$

If the system is in state a_i the i-th row vector represents the probabilities of all the possible outcomes of the next trial. It is, therefore, a probability vector. Hence,

THEOREM 2.6.2.1

The transition matrix **P** of a Markov chain is a stochastic matrix.
□

The entry p_{ij} of the transition matrix **P** of a Markov chain is the probability that the system changes from state a_i to state a_j in one step: $a_i \to a_j$. We want to know what is the probability, denoted by $p_{ij}^{(n)}$, that the system changes from state a_i to state a_j in exactly n steps:

$$a_i \to a_{k_1} \to \ldots \to a_{k_{n-1}} \to a_j$$

The next theorem answers this question. Here the $p_{ij}^{(n)}$ are arranged in a matrix $\mathbf{P}^{(n)}$ called the *n-step transition matrix.*

THEOREM 2.6.2.2

The n-step transition matrix of a Markov chain $\mathbf{P}$ is equal to the n-th power of $\mathbf{P}$. That is $\mathbf{P}^{(n)} = \mathbf{P}^n$.
□

Assume the probability that the system at time t is in state a_i is p_i. We denote these probabilities with the vector $\vec{p} = (p_1, p_2, ..., p_m)$ which is called the *probability distribution vector* of the system at time t. Let $\vec{p}^0 = \vec{p}\ p^{(0)} = (p_1^{(0)}, p_2^{(0)}, ..., p_m^{(0)})$ denote the *initial probability distribution* of the system at time $t = 0$. Let $\vec{p}^{(n)} = (p_1^{(n)}, p_2^{(n)}, ..., p_m^{(n)})$ denote the n-th *step probability distribution*. The following theorem applies.

THEOREM 2.6.2.3

Let $\mathbf{P}$ be the transition matrix of a Markov chain process. If $\vec{p} = (p_i)$ is the probability distribution at some time t, then $\vec{p}\mathbf{P}^n$ is the probability distribution of the system n steps later.
□

For example, $\vec{p}^{(1)} = \vec{p}^{(0)}\mathbf{P}$ and $\vec{p}^{(n)} = \vec{p}^{(0)}\mathbf{P}^n$.

THEOREM 2.6.2.4

If the transition matrix of a Markov chain is regular then $\lim_{t\to\infty} p(a_j) = \vec{t}_j$, where $\vec{t}$ is the fixed point of $\mathbf{P}$.