RANDOM EVOLUTIONS IN DISCRETE AND CONTINUOUS TIME

Joel E. COHEN

The Rockefeller University, 1230 York Avenue, New York, NY 10021, U.S.A.

This paper points out a connection between random evolutions and products of random matrices. This connection is useful in predicting the long-run growth rate of a single-type, continuously changing population in randomly varying environments using only observations at discrete points in time. A scalar Markov random evolution is specified by the $n \times n$ irreducible intensity matrix or infinitesimal generator $Q = (q_{ii})$ of a time-homogeneous Markov chain and by n finite real growth rates (scalars) s_i. The scalar Markov random evolution is the quantity $M^C(t) = \exp(\sum_{j=1}^{n} s_j g_j^C(t))$, where $g_j^C(t)$ is the occupancy times in state j up to time t. The scalar Markov product of random matrices induced by this scalar Markov random evolution is the quantity $M^D(t) = \exp(\sum_{j=1}^{n} s_j g_j^D(t))$, where $g_j^D(t)$ is the occupancy time in state j up to and including t of the *discrete*-time Markov chain with stochastic one-step transition matrix $P = e^Q$. We show that $\lim_{t \to \infty} (1/t)\mathbf{E}(\log M^D(t)) = \lim_{t \to \infty} (1/t)\mathbf{E}(\log M^C(t))$ but that in general $\lim_{t \to \infty} (1/t) \log \mathbf{E}(M^C(t)) \neq \lim_{t \to \infty} (1/t) \log \mathbf{E}(M^D(t))$. Thus the mean Malthusian parameter of population biologists is invariant with respect to the choice of continuous or discrete time, but the rate of growth of average population size is not. By contrast with a single-type population, in multitype populations whose growth is governed by non-commuting operators, the mean Malthusian parameter may be destined for a less prominent role as a measure of long-run growth.

Products of random matrices	Malthusian parameter
single-type populations	embedding
multi-type populations	

1. Introduction

Consider a homogeneous or single-type population such that, in a fixed environment, the magnitude of the population at time t, $M(t) \geq 0$, changes geometrically (in discrete time) or exponentially (in continuous time). Suppose that the population may experience, at any instant t, one of $n \geq 1$, n finite, different environments. A distinct (discrete or continuous) growth rate occurs in each environment. Suppose that the sequence of environments that a population experiences is determined by an n-state time-homogeneous Markov chain (in discrete time or in continuous time).

Certain products of random matrices and certain random evolutions may be viewed as descriptions in discrete and in continuous time, respectively, of such a stochastic growth process. We call this class of processes the "scalar Markovian

case" because the state of the population is specified by a scalar, which measures its total size, rather than by a vector, and because successive growth rates are Markovian, rather than determined by some other random process.

In both the discrete-time and the continuous-time scalar Markovian case, it is easy to evaluate each of two plausible measures of the long-run growth rate. It happens that only one of the two measures remains the same whether time is viewed continuously or discretely. If the scalar Markovian case were used as a model of some real process, it would be desirable for any measure of long-run growth rate to be invariant with respect to the choice of the description of time. Consequently the measure of growth rate that is invariant is to be preferred when the scalar Markov case is used as a model in science. This measure coincides with what population biologists call the mean Malthusian parameter.

We do not know of any previous systematic studies of the connection between products of random matrices and random evolutions.

2. The Markov scalar case

We define the Markov scalar case of a product of random matrices and of a random evolution.

As a special case of the products of random matrices considered by Furstenberg and Kesten [5], let X be a set of n finite positive real numbers. Suppose an n-state discrete-time Markov chain has state space X. If $\{X(t), t = 1, 2, \ldots\}$ is a sequence of members of X which occur as a sample path of this Markov chain,

$$M^D(t) = X(t)X(t-1) \cdots X(2)X(1) \tag{1}$$

is the scalar Markovian case of a product of random matrices. $M^D(t)$ may be thought of as the population size at time t. The superscript ^{D} refers to discrete time. Each additional instant of time multiplies the product $M^D(t)$ by a factor on the left.

As a special case of the random evolutions considered by Griego and Hersh [6, p. 407], we define a Markov scalar random evolution. Let $S = \{s_1, \ldots, s_n\}$ be any n finite real (positive, zero or negative) numbers. Let $v^C = \{v(t), t \geq 0\}$ be a continuous-time stationary Markov chain with state space S, time-homogeneous transition probabilities $p_{ij}(t) = \mathbf{P}[v(t) = s_j \,|\, v(0) = s_i]$, and infinitesimal matrix $Q = (q_{ij}) = (\mathrm{d}p_{ij}(0)/\mathrm{d}t)$. For $i \neq j$, $q_{ij} \geq 0$ and $q_{ii} = -\sum_{j \neq i} q_{ij}$. We assume Q is irreducible. $v(t)$ is the state the chain is in at time t. For any sample path of v^C, τ_j is the time of the jth jump and $N(t)$ is the number of jumps up to time t. Then

$$M^C(t) = \exp(s_{v(0)}\tau_1) \exp(s_{v(\tau_1)}(\tau_2 - \tau_1)) \cdots \exp(s_{v(N(t))}(t - \tau_{N(t)})) \tag{2}$$

is a Markov scalar random evolution. $M^C(t)$ may be thought of as the population size at time t. The superscript ^{C} indicates continuous time. Each new jump of the chain changes the current instantaneous rate of growth of the random evolution. Collecting

the factors in (2) which share a common value of s yields

$$M^C(t) = \exp\left(\sum_{j=1}^{n} s_j g_j^C(t) \right),\tag{3}$$

where $g_j^C(t)$ is the Markov chain's occupation time in the jth state up to time t.

Suppose now that a Markov scalar random evolution could be observed only at times $t = 1, 2, \ldots$ and that at each time t of observation it were possible to measure only the infinitesimal growth rate s_i in effect at that instant and the total magnitude $M^C(t)$ of the random evolution. Deprived of information about the course of events between $t-1$ and t, the observer attempts to construct a dynamic model of the trajectories of the infinitesimal generators $\{s_i\}$ that will explain the observed behavior of $M^C(t)$.

We shall suppose that the observer constructs the following *Markov scalar product of random matrices induced by the Markov scalar random evolutions*. Let $X = \{e^{s_1}, \ldots, e^{s_n}\}$. Let the stochastic process with state space X be an n-state stationary time-homogeneous Markov chain $v^D = \{v^D(t), t = 1, 2, \ldots\}$ with time-homogeneous one-step transition probability matrix $P = P(1) = e^Q$. Since Q is irreducible, each $q_{ii} < 0$ and $P(t) = (p_{ij}(t)) = e^{Qt}$ is positive (elementwise) for $t > 0$. Then

$$M^D(t) = \exp\left(\sum_{j=1}^{n} s_j g_j^D(t) \right),\tag{4}$$

where $g_j^D(t)$ is the number of times $v^D(h) = e^{s_j}$, $h = 1, \ldots, t$.

If π^T is the (unique) row-vector whose elements π_i (all positive) are the equilibrium probabilities of state i in the chain v^C, then $\pi^T Q = 0$ implies $\pi^T P = \pi^T$; the equilibrium probabilities of v^C and v^D are identical.

However, a sample path of the discrete Markov scalar case $M^D(t)$ differs from a sample path of the inducing continuous Markov scalar case $M^C(t)$. In $M^C(t)$, the operator $e^{s_i t}$, once it occurs, applies for a duration with a probability density function $q_{ii} e^{-q_{ii} t}$, $t \geq 0$. In $M^D(t)$, $e^{s_i t}$, once it occurs, applies for a duration with a probability density function $p_{ii}^{t-1}(1 - p_{ii})$, $t = 1, 2, \ldots$.

Now define, for these particular $M^C(t)$ and $M^D(t)$,

$$\log \lambda^C = \lim_{t \to \infty} \frac{1}{t} \mathbf{E}(\log M^C(t)),\tag{5}$$

$$\log \lambda^D = \lim_{t \to \infty} \frac{1}{t} \mathbf{E}(\log M^D(t)),\tag{6}$$

$$\log \mu^C = \lim_{t \to \infty} \frac{1}{t} \log \mathbf{E}(M^C(t)),\tag{7}$$

$$\log \mu^D = \lim_{t \to \infty} \frac{1}{t} \log \mathbf{E}(M^D(t)).\tag{8}$$

If these limits exist, they must be independent of the initial state of the discrete or continuous Markov chains because both chains are irreducible and aperiodic.

Substituting (3) into (5) and (4) into (6) gives

$$\log \lambda^C = \lim_{t \to \infty} \sum_{i=1}^{n} s_i \mathbf{E}\left(\frac{g_i^C(t)}{t}\right) = \sum_{i=1}^{n} s_i \pi_i, \tag{9}$$

$$\log \lambda^D = \lim_{t \to \infty} \sum_{i=1}^{n} s_i \mathbf{E}\left(\frac{g_i^D(t)}{t}\right) = \log \lambda^C. \tag{10}$$

Thus a hypothetical observer restricted to discrete times of observation who constructs the scalar Markov product of random matrices induced by a scalar Markov random evolution will predict exactly the growth rate λ^C by calculating the geometric mean λ^D on the basis of his construction. The log of the geometric mean $\sum_{i=1}^{n} s_i \pi_i$, familiar to population biologists [10] as the "mean Malthusian parameter", finds a new interpretation in the scalar Markovian case through (5), (6), (9) and (10).

To calculate μ^C from (7), let $\mathbf{E}(\exp[\sum s_j g_j^C(t)] | v(0) = i) = u_i(t)$, and let $S = \mathrm{diag}\{s_1, \ldots, s_n\}$.

The Feynman–Kac formula for random evolutions with commuting families of operators developed by Griego and Hersh [6] asserts that if $A = Q + S$, i.e., $a_{ii} = q_{ii} + s_i$, $a_{ij} = q_{ij}$ $(i \neq j)$, and if $u(t)$ is the n-vector with ith element $u_i(t)$, then $u(t)$ is the unique solution of

$$\mathrm{d}u/\mathrm{d}t = Au, \qquad u_i(0) = 1. \tag{11}$$

But it is well known that the unique solution of (11) is

$$u(t) = \mathrm{e}^{At} u(0), \tag{12}$$

i.e., $u_i(t) = i$th row sum of e^{At}. We adopt the norm, for any complex matrix $M = (m_{ij})$, $\|M\| = \max_i \sum_{j=1}^{n} |m_{ij}|$. Then $\max_i u_i(t) = \|\mathrm{e}^{At}\|$. Because Q is irreducible, A is irreducible. Hence so is e^A. By the Jordan canonical form of A, the spectral radius $r(A)$ of A satisfies $r(A) = \log r(\mathrm{e}^A)$. But $r(\mathrm{e}^A) = \lim_{t \to \infty} \|\mathrm{e}^{At}\|^{1/t}$ for any matrix norm, so

$$r(A) = \log \lim_{t \to \infty} \|\mathrm{e}^{At}\|^{1/t} = \lim_{t \to \infty} \frac{1}{t} \log \max_i \mathbf{E}\left(\exp\left[\sum_j s_j g_j(t)\right] \Big| v(0) = i\right) = \log \mu^C,$$

since μ^C is independent of $v(0)$. Collecting equalities, we have $\mu^C = r(\mathrm{e}^{Q+S})$. (This argument also leads readily to the conclusion that if A is a non-negative $n \times n$ matrix and V is a real $n \times n$ diagonal matrix, then $r(A + V)$ is a convex function of V. The additional requirement in [4, Theorem 3] that V be non-negative is evidently irrelevant.)

To compute μ^D, we apply to the scalar Markov product of random matrices a discrete-time Feynman–Kac formula [1], which was not recognized as such when first reported. Since $P = (p_{ij}) = P(1) = \mathrm{e}^Q$, $\mathrm{e}^Q \mathrm{e}^S$ has i,jth element $p_{ij} \mathrm{e}^{s_j}$. Then by [1],

$$\lim_{t \to \infty} \mathbf{E}(M^D(t)) / [r(\mathrm{e}^Q \mathrm{e}^S)]^t = c > 0,$$

where c is a constant independent of time. Taking $(1/t)$ log of both sides and moving the limit outside gives the recipe $\mu^D = r(e^Q e^S)$.

We wish to compare $\mu^C = r(e^{Q+S})$ with $\mu^D = r(e^Q e^S)$. A simple numerical example in the case $n = 2$ shows that, in general, one may have $\mu^C < \mu^D$. When $n = 2$, e^Q is easily written as an explicit function of Q and the spectral radius is the solution of a quadratic equation, so $r(e^{Q+S}) = e^{r(Q+S)}$ is easily found. Let

$$Q = \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix}, \qquad S = \begin{pmatrix} 4 & 0 \\ 0 & 3 \end{pmatrix}.$$

Then $r(e^{Q+S}) = 41.76 < r(e^Q e^S) = 43.41$. Thus μ^D, the rate of growth of the average scalar Markov product of random matrices induced by a scalar Markov random evolution, may in general be strictly greater than μ^C, the rate of growth of the average scalar Markov random evolution.

3. Further problems: matrix operators; embedding

A natural next case to consider is the matrix Markovian case. Suppose the positive scalars s_i are replaced by elementwise positive non-commuting $k \times k$ matrices M_i, $k > 1$ (possibly subject to suitable auxiliary conditions), and the induced product of random matrices has factors $X(t)$ chosen from the non-commuting set $X = \{e^{M_1}, \ldots e^{M_n}\}$ by a Markov chain with $P = e^Q$. Since $M^D(t)$ and $M^C(t)$ are now matrix rather than scalar functions of time, the definitions (5)–(8) must be modified, say, by replacing $M(t)$ with $\|M(t)\|$. The question is: does the equality $\lambda^C = \lambda^D$ still hold in general or in any interesting special cases? Exact equations for λ^D are known [2] and have been solved numerically [3] in particular cases. The problem is to compute λ^C.

At least when X is a finite "ergodic set" [7] of non-negative matrices, almost every sample path of a matrix Markov product of random matrices has a long-run rate of growth given by λ^D[2,5]. If r_i is the log of the spectral radius of the ith matrix in X (population biologists call r_i the Malthusian parameter of a k-type population in environment i), in general $\sum_i \pi_i r_i \neq \lambda^D$ [3]. This inequality is a significant difference between single-type and multitype populations. The mean Malthusian parameter $\sum_i \pi_i r_i$ is not interpretable as a measure of growth for multitype populations.

Another interesting, and probably more difficult, problem is the embedding problem: given a Markovian product of a random matrices, when could it have been induced (in the sense defined) by a random evolution, and when could it have been induced by only one random evolution? Half of the question is: given a stochastic $n \times n$ positive matrix P, when does there exist an intensity matrix Q such that $P = e^Q$, and when is Q unique? This is the embedding problem for Markov chains [9]. Kendall obtained a complete solution for $n = 2$, but the general case is recalcitrant. The other half of the question is: when can the matrices in X be written (uniquely) in the form e^{M_i}, $i = 1, \ldots, n$ with infinitesimal generators M_i of some specified form?

4. Population biology and life histories

Population biologists have modeled the life histories of various organisms. They have argued that evolution by natural selection would favor the life history or histories associated with the largest value of some characteristic of a population, such as growth rate or time to extinction.

In one example [11, pp. 165–168; 8], a single-type population of annual plants is compared with a two-type population of perennial plants. Under certain assumptions about the environment, one index of population growth favors an annual life history while another index of population growth favors a perennial life history. The questions arises, "Which index of growth (if any) should the biologist choose as a reliable guide to the life history that will evolve in an environment like the one assumed?"

Rational argument cannot answer that empirical question. But if one desired an index of growth in population size that is invariant under equivalent discrete and continuous descriptions of time, (6) would be preferred for a single-type population. This index λ^D is described as Cohen's measure of growth in [11] and as the geometric mean of the growth rate in [8]. For multitype populations, see [3].

Acknowledgment

References

[1] J.E. Cohen, Ergodicity of age structure in populations with Markovian vital rates. III. Finite-state moments and growth rates; illustration, Advances in Appl. Probability 9 (1977) 462–475.

[2] J.E. Cohen, Long-run growth rates of discrete multiplicative processes in Markovian environments, J. Math. Anal. Appl. 69 (1979) 243–251.

[3] J.E. Cohen, Comparative statics and stochastic dynamics of age-structured populations, Theoret. Population Biology, 16 (1979) 159–171.

[4] J.E. Cohen, Random evolutions and the spectral radius of a non-negative matrix, Math. Proc. Cambridge Philos. Soc., 86 (1979) 345–350.

[5] H. Furstenberg and H. Kesten, Products of random matrices, Ann. Math. Statist. 31(2) (1960) 457–469.

[6] R. Griego and R. Hersh, Theory of random evolutions with applications to partial differential equations, Trans. Amer. Math. Soc. 156 (1971) 405–418.

[7] J. Hajnal, On products of non-negative matrices, Math. Proc. Cambridge Philos. Soc. 79 (1976) 521–530.

[8] A. Hastings and H. Caswell, Environmental variability and the evolution of life history strategies, Proc. Nat. Acad. Sci. U.S.A., to appear.

[9] J.F.C. Kingman, The imbedding problem for finite Markov chains, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 1 (1962) 14–24.

[10] R.C. Lewontin and D. Cohen, On population growth in a randomly varying environment, Proc. Nat. Acad. Sci. U.S.A. 62(4) (1969) 1056–1060.

[11] R.H. MacArthur, Geographical Ecology: Patterns in the Distribution of Species (Harper and Row, New York, 1972).