SIAM REVIEW
Vol. 30, No. 1, March 1988

002

SUBADDITIVITY, GENERALIZED PRODUCTS OF RANDOM MATRICES AND OPERATIONS RESEARCH*

JOEL E. COHEN†

Abstract. An elementary theorem on subadditive sequences provides the key to a far-reaching theory of subadditive processes. One important instance of this theory is the limit theory for products of stationary random matrices. This paper shows that the subadditive inequality that governs the log of the norm of ordinary matrix products also governs other functions of several generalized matrix products. These generalized matrix products are used to calculate minimal cost transportation routes, schedules in manufacturing and minimal and maximal probabilities of multistage processes. The application of subadditive ergodic theory to generalized products of stationary random matrices yields new information about the limiting behavior of generalized products. Exact calculations of the asymptotic behavior are possible in some examples.

Key words. subadditive ergodic theory, products of random matrices, strong laws, generalized spectral radius, Furstenberg–Kesten limit, Lyapunov exponents, scheduling, transportation, reliability, asymptotic extreme value theory

AMS(MOS) subject classifications. 60K30, 90B15

1. Introduction. This paper traces a path from an elementary theorem of analysis to some problems of operations research, by way of some generalizations of matrix multiplication. Almost all the pieces of the path are elementary and well known. But different pieces are known to different people. The connection among the pieces seems new.

As we trace this path, known theorems will be labeled by capital letters (e.g., Theorem A) and new theorems by numbers (e.g., Theorem 1). We assume familiarity with basic linear algebra (e.g., Lancaster and Tismenetsky (1985)) and basic probability theory (e.g., Breiman (1968)), and we minimize measure-theoretic aspects not essential to grasping the meaning of statements.

For the reader who is already familiar with subadditive ergodic theory and products of random matrices, a reasonable place to continue reading would be §4.

2. Subadditivity and the subadditive ergodic theorem. Pólya and Szegö (1976, p. 23, Ex. I.3.98, p. 198) give the basic theorem of subadditivity.

THEOREM A. *Let* $\{a_1, a_2, a_3, \cdots\}$ *be a sequence of real numbers such that*

$$a_{m+n} \leqq a_m + a_n, \qquad m, n = 1, 2, 3, \cdots. \tag{1}$$

Then the sequence $\{a_n/n, n = 1, 2, 3, \cdots\}$ *either converges to its lower bound* $\gamma\{a_n\} = \inf_{n\geqq 1} a_n/n$ *or diverges properly to* $-\infty$.

To appreciate Theorem A, suppose that the assumption (1) of subadditivity were replaced by the hypothesis of additivity: $a_{m+n} = a_m + a_n$, for $m, n = 1, 2, \cdots$. By immediate induction, $a_n = na_1$ or $a_n/n = a_1$ for all n. In an additive sequence, the nth term a_n is exactly proportional to n. Theorem A asserts that under the subadditive bound (1) on the growth of the sequence, if a_n/n does not fall to $-\infty$, then, as in

*Received by the editors December 30, 1985; accepted for publication February 11, 1987. The Inter-Academy Exchange Program of the Romanian Academy, Bucharest, Romania, and the U.S. National Academy of Sciences made possible the author's visit to the Centre of Mathematical Statistics. This work was supported in part by National Science Foundation grant BSR84-07461 and a Fellowship from the John D. and Catherine T. MacArthur Foundation.

† The Rockefeller University, New York, New York 10021-6399.

the additive case, a_n/n has a finite limit and the sequence $\{a_n\}$ is asymptotically proportional to n.

Proof of Theorem A (Pólya and Szegö (1976)). Let $\gamma = \inf_{n \geqq 1} a_n/n$. If $\gamma = -\infty$, there is nothing to prove. If $\gamma > -\infty$, pick any $\varepsilon > 0$ and find a fixed m such that $a_m/m < \gamma + \varepsilon$. Define $a_0 = 0$. Since any integer n can be written $n = qm + r$ with r an integer such that $0 \leqq r \leqq m - 1$, subadditivity implies

$$a_n = a_{qm+r} \leqq a_m + \cdots (q \text{ times}) + a_m + a_r = qa_m + a_r.$$

Hence

$$\gamma \leqq a_n/n \leqq (qa_m + a_r)/n = qa_m/n + a_r/n$$
$$= (a_m/m)(qm/n) + a_r/n < (\gamma + \varepsilon)(qm/n) + a_r/n.$$

Now let $n \uparrow \infty$. Then $qm/n \to 1$, $a_r/n \to 0$. Since $\varepsilon > 0$ was arbitrary, $a_n/n \to \gamma$. □

Kingman (1968) (see also (1973), (1976)) developed a far-reaching probabilistic application of this theorem. Let T (think of time) be the set of nonnegative integers. Let $\mathbf{x}$ be a family $\{x_{st} | s, t \in T, s < t\}$ of random variables x_{st} which may be thought of as describing certain random events that occur after time s up to and including time t.

A *subadditive process* is defined as a family $\mathbf{x}$ such that

(S1) $x_{su} \leqq x_{st} + x_{tu}$ for all $s < t < u$, $s, t, u \in T$;

(S2) All joint distributions of the shifted family $\mathbf{x}' = \{x_{s+1,t+1} | s, t \in T, s < t\}$ are the same as those of $\mathbf{x} = \{x_{st} | s, t \in T, s < t\}$;

(S3) The expectation $g_t = \mathrm{E}(x_{0t})$ exists and satisfies $g_t \geqq -ct$, for some finite constant c and every $t \geqq 1$.

The stationarity condition (S2) implies that, for any $s < t$, $\mathrm{E}(x_{st}) = \mathrm{E}(x_{s+1,t+1})$, i.e., $\mathrm{E}(x_{st})$ depends only on $t - s$. Hence $\mathrm{E}(x_{st}) = \mathrm{E}(x_{0,t-s}) = g_{t-s}$. Averaging (S1) gives

$$g_{u-s} \leqq g_{t-s} + g_{u-t}, \qquad s < t < u,$$

or

$$g_{m+n} \leqq g_m + g_n, \qquad m, n \geqq 1.$$

Since (S3) gives $g_t/t \geqq -c$, for all $t \in T$, Theorem A implies $\lim_{t \uparrow \infty} g_t/t$ exists and equals the finite number

$$\gamma = \gamma(\mathbf{x}) = \inf_{t \geqq 1} g_t/t.$$

This is a major part, but not all, of Kingman's subadditive ergodic theorem.

THEOREM B. *If* $\mathbf{x}$ *is a subadditive process, then with probability one there exists a limiting finite-valued random variable*

$$\xi = \lim_{t \uparrow \infty} x_{0t}/t$$

which has a finite expectation, and

$$\mathrm{E}(\xi) = \gamma(\mathbf{x}).$$

The additional point that we have not proved in Kingman's theorem is that x_{0t}/t converges to a finite but possibly random limit ξ in almost every sample path of the process, not merely on the average. The finite average $\mathrm{E}(\xi)$ of the limits of x_{0t}/t equals the limit $\gamma(\mathbf{x})$ of the averages g_t/t.

In many applications of the subadditive ergodic theorem, it can be proved that the limiting random variable is degenerate, i.e., takes only a single value with probability one. This value must be $\gamma(\mathbf{x})$, so in this case,

$$\lim_{t\uparrow\infty} x_{0t}/t=\gamma(\mathbf{x}) \quad \text{almost surely.}$$

3. Products of random matrices. The subadditive ergodic theorem has remarkable implications for products of random matrices.

Let d be a fixed positive integer, $1<d<\infty$. (Everything that follows is true, but trivial, for $d=1$ as well.) Let $\mathbb{C}^{d\times d}$, $\mathbb{R}^{d\times d}$, $\mathbb{R}_+^{d\times d}$ and $\mathbb{R}_{++}^{d\times d}$ be the sets of $d\times d$ matrices over, respectively, the complex numbers, the real numbers, the nonnegative reals and the positive reals.

A matrix norm $\|A\|$ is a real-valued function of matrices $A\in\mathbb{C}^{d\times d}$ such that

(M1) $\|A\|\geqq 0$ and $\|A\|=0$ if and only if $A=0$;

(M2) For any $c\in\mathbb{C}$, $\|cA\|=|c|\cdot\|A\|$;

(M3) $\|A+B\|\leqq\|A\|+\|B\|$ for $A, B\in\mathbb{C}^{d\times d}$;

(M4) $\|AB\|\leqq\|A\|\cdot\|B\|$ for $A, B\in\mathbb{C}^{d\times d}$.

The crude resemblance between the triangle inequality (M3) and subadditivity (1) suggests an obvious application of the subadditive ergodic theorem. Theorem 1 illustrates the use of the subadditive ergodic theorem under very easy circumstances.

THEOREM 1. *Let* $\{A_t\}_{t\in T}$ *be a stationary sequence of random matrices from* $\mathbb{C}^{d\times d}$ *and let*

$$S_{st}=A_{s+1}+\cdots+A_{t-1}+A_t, \qquad s,t\in T,\quad s<t.$$

If $\mathrm{E}\|A_1\|<\infty$, *then*

$$\lim_{t\uparrow\infty}\|S_{0t}\|/t=\xi$$

exists with probability one, is finite, and has an expectation $\mathrm{E}(\xi)$ *that satisfies*

$$\mathrm{E}(\xi)=\lim_{t\uparrow\infty}\mathrm{E}\{\|S_{0t}\|\}/t\equiv\gamma_+\{A_t\}.$$

If, in addition, the random matrices $\{A_t\}$ *are independently and identically distributed (i.i.d.), then*

$$\xi=\gamma_+\{A_t\} \quad \textit{with probability one.}$$

Proof. Let $x_{st}=\|S_{st}\|$. Is $\{x_{st}\}=\mathbf{x}$ a subadditive process? (S1) For $s<t<u$, $s,t,u\in T$, $x_{su}=\|S_{su}\|=\|S_{st}+S_{tu}\|\leqq$ (by M3) $\|S_{st}\|+\|S_{tu}\|=x_{st}+x_{tu}$. (S2) Since $\{A_t\}$ is stationary, $\mathbf{x}'$ and $\mathbf{x}$ have the same joint distributions. (S3) $\mathrm{E}(x_{0t})=\mathrm{E}\|S_{0t}\|\leqq\mathrm{E}(t\|A_1\|)=t\mathrm{E}\|A_1\|<\infty$. Hence $\mathrm{E}(x_{0t})\equiv g_t$ exists. Moreover $g_t\geqq 0$ by (M1). Thus $\{x_{st}\}$ is a subadditive process. The subadditive ergodic theorem then applies.

If $\{A_t\}$ are i.i.d., then the only σ-field of events defined in terms of $\mathbf{x}$ and invariant under the shift $\mathbf{x}\rightarrow\mathbf{x}'$ is the trivial σ-field, so $\xi=\gamma_+\{A_t\}$ almost surely. □

A much more important instance of the subadditive ergodic theorem follows from the trivial observation that (M4) implies $\log\|AB\|\leqq\log\|A\|+\log\|B\|$. First consider the simplest possible application of this inequality.

Fix $A\in\mathbb{C}^{d\times d}$ and let $x_{st}=\log\|A^{t-s}\|$, $s<t$, $s,t\in T$. The family $\mathbf{x}$ of degenerate random variables $\{x_{st}\}$ constitutes a subadditive process with a degenerate limit random variable $\xi=\lim_{t\uparrow\infty}t^{-1}\log\|A^t\|$, which equals almost surely

$\gamma\{\mathbf{x}\} = \inf_{t\geqq 1} t^{-1} \log \|A^t\| = \log \inf_{t\geqq 1} \|A^t\|^{1/t} \equiv \log \rho(A)$, where $\rho(A)$ is known as the spectral radius of A. The spectral radius is the maximum of the moduli of the eigenvalues of A and measures how fast, on the average, $\|A^t\|$ grows asymptotically with each additional power of A.

In 1960, even before there was any general subadditive ergodic theory, Furstenberg and Kesten discovered a very important generalization of the spectral radius of one matrix. They discovered that products of random matrices generated by a stationary process also have, under reasonable conditions, an asymptotic growth rate. This growth rate corresponds exactly to the logarithm of the spectral radius when all the random matrices degenerate to a single matrix.

THEOREM C. *Let* $\|\cdot\|$ *be a nonnegative real-valued function of matrices that satisfies* (M4). *Let* $\{A_t\}_{t\in T}$ *be a stationary sequence of random matrices from* $\mathbb{C}^{d\times d}$ *and let*

$$P_{st} = A_{s+1}A_{s+2}\cdots A_{t-1}A_t, \qquad s, t\in T, \quad s<t.$$

If $\mathrm{E}\{\max(0, \log \|A_1\|)\} < \infty$, *then*

$$\lim_{t\uparrow\infty} t^{-1} \log \|P_{0t}\| = \xi$$

exists with probability one, $-\infty \leqq \xi < \infty$, *and has an expectation* $\mathrm{E}(\xi)$ *that satisfies*

$$-\infty \leqq \mathrm{E}(\xi) = \lim_{t\uparrow\infty} t^{-1}\mathrm{E}(\log \|P_{0t}\|) \equiv \gamma_{+\cdot\times}\{A_t\} < \infty.$$

If $\{A_t\}$ *are i.i.d., then* $\xi = \gamma_{+\cdot\times}\{A_t\}$ *almost surely.*

Let $x_{st} = \log \|P_{st}\|$. Then $\mathbf{x} = \{x_{st} | s, t\in T, s<t\}$ satisfies (S1) and (S2) but not (S3); hence there is no guarantee that $\mathrm{E}(\xi) > -\infty$.

Furstenberg and Kesten (1960) and Kingman (1973, pp. 891–892) also considered the special case when all elements of every matrix are positive, or when there exists $N\in T$ such that every $P_{s,s+N}$ has all elements positive.

THEOREM D. *Let* $\{A_t\}_{t\in T}$ *be a stationary sequence of random matrices from* $\mathbb{R}_{++}^{d\times d}$. *Define* P_{st} *as in Theorem* C. *If* $-\infty < \mathrm{E}\{\log (A_t)_{ij}\} < \infty$, *for all* $t\in T$, $1 \leqq i, j \leqq d$, *then*

$$\lim_{t\uparrow\infty} t^{-1} \log (P_{0t})_{ij} = \xi$$

exists with probability one, is independent of i and j, and has a finite mean

$$\mathrm{E}(\xi) = \lim_{t\uparrow\infty} t^{-1}\mathrm{E}\{\log (P_{0t})_{11}\} \equiv \gamma_{+\cdot\times}\{A_t\}.$$

If $\{A_t\}$ *are i.i.d., then* $\xi = \gamma_{+\cdot\times}\{A_t\}$ *almost surely.*

Kingman's proof (1973, pp. 891–892) of Theorem D generalizes easily to variants of ordinary matrix multiplication. A proof that is essentially Kingman's in the context of generalized matrix multiplication is given below.

Theorem D is a beautiful generalization of the Perron–Frobenius Theorem on positive matrices. Let $A\in\mathbb{R}_{++}^{d\times d}$. Then A^t corresponds to P_{0t} for random matrix products. The Perron–Frobenius Theorem asserts that $\lim_{t\uparrow\infty} t^{-1} \log (A^t)_{ij}$ exists, is independent of i, j, and equals $\log \rho(A)$.

The key quantity in Theorems C and D is $\gamma = \gamma_{+\cdot\times}\{A_t\}$, the asymptotic growth rate. If $\gamma > 0$, then asymptotically $\|P_{0t}\|$ grows exponentially with t; if $-\infty < \gamma < 0$, $\|P_{0t}\|$ declines exponentially with t. The first information about the sign of γ appeared in work of Furstenberg (1963). A decade later, Kingman (1973, p. 897) wrote, in introducing a list of open problems: "Pride of place among the unsolved problems of subadditive ergodic theory must go to the calculation of the constant γ (which in the presence of a zero-one law is the same as the limit ξ). In none of the applications described here is there an obvious mechanism for obtaining an exact numerical value, and indeed this usually seems to be a problem of some depth."

It is reasonable to expect the computation of γ to be at least as difficult as the computation of the spectral radius $\rho(A)$ of a fixed matrix A. In general, $\rho(A)$ can only be calculated numerically, but for some special cases of A, simple formulas are possible. For example, if $A \in \mathbb{R}_{++}^{d\times d}$ and the sum of the elements in every row is the same and equal to $c>0$, then $\rho(A)=c$.

For products of random matrices, one simple case has been analyzed (Cohen and Newman (1984)). Let $\{A_t\}$ be i.i.d. matrices from $\mathbb{R}^{d\times d}$ in which all d^2 elements of each A_t are i.i.d. normal random variables with mean 0 and variance 1. Then, in the sense of Theorem C, almost surely

$$\xi=(\tfrac{1}{2})\log 2+(\tfrac{1}{2})\Psi(d/2)$$

where Ψ is the digamma function.

In summary of this section, a limit theory describes the growth of subadditive functions of products of random matrices. These functions may pertain to overall matrix size (like the log of the matrix norm), as in Theorem C, to individual matrix elements, as in Theorem D, or to subtler aspects of matrix structure (like the log of the coefficient of ergodicity or convergence norm (Kingman (1976, p. 197)) or the log of Birkhoff's contraction coefficient (Hajnal (1976))). Sometimes the limiting growth rate can be computed explicitly.

Since the pioneering papers of Bellman (1954) and Furstenberg and Kesten (1960), the theory of products of random matrices has developed enormously. The theory has found applications in number theory, ergodic theory, statistics, computer science, statistical physics, quantum mechanics, ecology and demography. A recent expository collection (Cohen, Kesten and Newman (1986)) samples recent theory and applications. Bougerol and Lacroix (1985) give a coherent account of the theory of products of random nonsingular matrices. I now turn to an apparently new application of subadditive ergodic theory to generalized products of random matrices.

4. Generalized matrix multiplication. Numerous generalizations of matrix multiplication have been discovered, independently, to be of mathematical and practical interest. Without making any attempt at a history, we will present and interpret several of these generalizations. Recent monographs devoted entirely to generalizations of matrix multiplication include those of Cuninghame-Green (1979), Hammer and Rudeanu (1968) and Kim (1982).

Let F denote the set of numbers (e.g., $\mathbb{C}$, $\mathbb{R}_{++}$, $\{0, 1\}$, etc.) that appear as elements of the set of matrices being considered. Let f and g be functions from $F\times F$ into F. For simplicity, assume f to be associative.

Define $f\cdot g$ to be the function from $F^{d\times d'}\times F^{d'\times d''}$ into $F^{d\times d''}$ given, for $(a_{ij})=A\in F^{d\times d'}$, $(b_{jk})=B\in F^{d'\times d''}$, by

$$(Af\cdot gB)_{ik}=(a_{i1}gb_{1k})f(a_{i2}gb_{2k})f\cdots f(a_{id'}gb_{d'k}),$$
$$i=1,\cdots,d,\qquad k=1,\cdots,d''.$$

In this notation, ordinary matrix multiplication is written $A+\cdot\times B$. The notation $f\cdot g$ is used in the contemporary programming language APL (International Business Machines (1983)). This generalized matrix multiplication appeared as an element of APL at least as early as 1962 (Iverson (1962, pp. 23–25)). Boolean special cases of generalized matrix multiplication were studied at least a decade before that (Lunc (1950), (1952); Cetlin (1952)).

To assure that $f \cdot g$-multiplication is associative, i.e., that

$$(Af \cdot gB)f \cdot gC = Af \cdot g(Bf \cdot gC)$$

for any matrices over F for which the operations are defined, it suffices to assume that both f and g are associative (not only f, as assumed in defining $f \cdot g$), f is commutative, and g distributes over f, i.e.,

$$(afb)gc = (agc)f(bgc) \quad \text{for all } a, b, c, \in F.$$

For the remainder of this paper, we suppose that f and g are limited to four possible binary functions, $+$, $\times$, $\llcorner$, and $\ulcorner$. Here $+$ is ordinary addition over F, $\times$ is ordinary multiplication over F, and $\llcorner$ (the notation called "floor" in APL) and $\ulcorner$ (the notation called "ceiling" in APL) represent, respectively, min and max. $\llcorner$ and $\ulcorner$ are defined only over $\mathbb{R}$ or over subsets of $\mathbb{R}$, not over $\mathbb{C}$. I will denote by **O** the set $\{+, \times, \llcorner, \ulcorner\}$. (In APL, the set of possible binary functions in generalized matrix multiplication is much larger than **O**. For the concrete applications of interest here, **O** suffices.)

Four possibilities for f and four possibilities for g define 16 possibilities for $f \cdot g$. Not all of these $f \cdot g$-multiplications are associative. We will now try to show why five of these generalized matrix multiplications are of interest by giving concrete interpretations of them (Moisil (1960); Cuninghame-Green (1979)). In these examples, $f \cdot g$-multiplication is associative.

Ace Distributors serves d sites (stores and factories), with labels i, j or $k = 1, 2, \cdots, d$. On day 1, the cost of sending a truck from site i to site j is a_{ij}. On day 2, the cost of sending a truck (not necessarily the same truck) from site j to site k is b_{jk}. Since certain trips may be heavily subsidized, e.g., by the government, there is no reason to exclude the possibility that a cost a_{ij} or b_{jk} may be negative. With $A = (a_{ij})$, $B = (b_{jk})$, the minimum cost of sending a sequence of trucks from site i to site k, starting on day 1 and ending on day 2, is $(A\llcorner \cdot + B)_{ik}$, so $A \llcorner \cdot + B$ is the matrix of minimum costs of two-stage trips. The minimum cost of a t-stage trip from site i to site k is $(A_1 \llcorner \cdot + \cdots \llcorner \cdot + A_t)_{ik}$ if A_t is the cost matrix for stage t. The matrix of average (per stage) minimum costs for t-stage trips is t^{-1} $(A_1 \llcorner \cdot + \cdots \llcorner \cdot + A_t)$. Similarly, the matrix of average (per stage) maximum costs for t-stage trips is t^{-1} $(A_1 \ulcorner \cdot + \cdots \ulcorner \cdot + A_t)$.

Different trucks have different capacities. If $(A_t)_{ij}$ is the maximum weight Ace can ship from site i to site j on day $t \in T$, then the maximum weight Ace can ship from site i via site j to site k starting on day 1 and ending on day 2 is $(A_1)_{ij} \llcorner (A_2)_{jk}$. Hence the maximum weight Ace can ship from site i to site k via any intermediate site, starting on day 1 and ending on day 2, is the previous amount maximized over j, or $(A_1 \ulcorner \cdot \llcorner A_2)_{ik}$. Thus $A_1 \ulcorner \cdot \llcorner \cdots \ulcorner \cdot \llcorner A_t$ is the matrix of shipping capacity over t-stage trips and $t^{-1}(A_1 \ulcorner \cdot \llcorner \cdots \ulcorner \cdot \llcorner A_t)$ is the matrix of average (per stage) capacity in t-stage trips. A negative capacity a_{ij} could be interpreted as a guaranteed loss of goods in transit from i to j. When restricted to matrices with Boolean elements $\{0, 1\}$, $\ulcorner \cdot \llcorner$ is identical to Boolean multiplication of Boolean matrices.

Ace's trucks occasionally break down. If $(A_t)_{ij}$ is the probability that the shipment from site i to site j on day t can be completed successfully, and if shipments succeed or fail independently on different days, then the probability of a successful shipment from site i via site j to site k starting on day 1 and ending on day 2 is $(A_1)_{ij}(A_2)_{jk}$. Therefore the maximum probability of success of a shipment from site i

to site k via any intermediate site, starting on day 1 and ending on day 2, is $(A_1 \ulcorner \cdot \times A_2)_{ik}$ and the minimum probability of success is $(A_1 \llcorner \cdot \times A_2)_{ik}$. Thus $A_1 \ulcorner \cdot \times \cdots \ulcorner \cdot \times A_t$ is the matrix of maximum probabilities of successful shipment over t-stage trips and $[(A_1 \ulcorner \cdot \times \cdots \ulcorner \cdot \times A_t)_{ik}]^{1/t}$ is the geometric mean maximum probability of successful shipment over t-stage trips from site i to site k. $\times$ distributes over $\ulcorner$ on $\mathbb{R}_+$ but not on $\mathbb{R}$; hence, as the context implies, A_t are limited to $\mathbb{R}_+^{d\times d}$.

Ace distributes for Best Manufacturing. Best uses d machines. Each machine has a cycle, an operation that lasts for a variable period of time. After machine j begins its $(t-1)$st cycle, it must wait for some or all of the machines, including itself, to operate (in parallel) before it can begin its tth cycle. Let $y(t)$ be the vector in which the jth element $y_j(t)$ is the time when the jth machine begins its tth cycle, $j = 1, \cdots, d$, $t \in T$. Let $(A_t)_{ij}$ be the duration of operation of machine i when operating between the start of cycle $t-1$ and the start of cycle t of machine j. Thus $(A_t)_{ii}$ is the duration of machine i's operation during its $(t-1)$st cycle. Let $(A_t)_{ij} = -\infty$ if machine i need not operate between cycles of machine j. Then $y_i(t)$ is the maximum over j of $(A_t)_{ij} + y_j(t-1)$ (letting $-\infty + a = -\infty$ for any real number a), and $y(t) = A_t \ulcorner \cdot + y(t-1)$. By induction, $y(t) = A_t \ulcorner \cdot + \cdots \ulcorner \cdot + A_1 \ulcorner \cdot + y(0)$. Then the average interval between cycles of each machine is given by the vector $(y(t) - y(0))/t$.

These little stories demonstrate the practical interest of various generalizations of matrix multiplication. There are many other applications, for example, to computing the distance matrix of a graph and the shortest path matrix in a general network. It is an amusing exercise to pick any of the remaining multiplications $f \cdot g$ with $f, g \in \mathbf{O}$, and attempt to interpret it concretely.

5. Subadditive ergodic theory of generalized matrix products. When $f \cdot g$-multiplication $(f, g \in \mathbf{O} = \{\ulcorner, \llcorner, +, \times\})$ is associative for $d \times d$ matrices with elements from $F \subseteq \mathbb{R}$, define

$$P_{st}^{f\cdot g} = A_{s+1} f \cdot g A_{s+2} f \cdot g \cdots f \cdot g A_{t-1} f \cdot g A_t, \qquad s, t \in T, \quad s < t.$$

When the context specifies $f \cdot g$, we abbreviate $P_{st}^{f\cdot g}$ to P_{st}. (Thus, in Theorems C and D, $P_{st} = P_{st}^{+\cdot\times}$.)

For any matrix $A \in \mathbb{C}^{d\times d}$, $A = (a_{ij})$, define

$$\|A\|_1 = \max_i \sum_j |a_{ij}|.$$

Note that $|a_{11}| \leqq \|A\|_1$.

The next theorem is a variation on the theme of Theorem D, and its proof is Kingman's (1973, pp. 891–892) proof of Theorem D except for minor modifications.

THEOREM 2. *Let $\{A_t\}_{t\in T}$ be a stationary sequence of random matrices from $\mathbb{R}_{++}^{d\times d}$. If $-\infty < \mathrm{E}\{\log (A_t)_{ij}\} < \infty$ for all $t \in T$, $1 \leqq i, j \leqq d$, then*

$$\lim_{t\uparrow\infty} t^{-1} \log (P_{0t}^{\ulcorner\cdot\times})_{ij} = \xi$$

exists with probability one, is independent of i and j, and has a finite mean

$$\mathrm{E}(\xi) = \lim_{t\uparrow\infty} t^{-1}\mathrm{E}\{\log (P_{0t}^{\ulcorner\cdot\times})_{11}\} \equiv \gamma_{\ulcorner\cdot\times}\{A_t\}.$$

If $\{A_t\}$ are i.i.d., then $\xi = \gamma_{\ulcorner\cdot\times}\{A_t\}$ almost surely.

Proof. For the balance of this proof, fix $f \cdot g = \ulcorner \cdot \times$. This multiplication is associative on $\mathbb{R}_{++}^{d\times d}$. Hence, for $s < t < u$, $s, t, u \in T$, $P_{su} = P_{st} \ulcorner \cdot \times P_{tu}$ and

$$(P_{su})_{11} = \max_j (P_{st})_{1j}(P_{tu})_{j1} \geqq (P_{st})_{11}(P_{tu})_{11}.$$

Then

$$x_{st} \equiv -\log (P_{st})_{11}, \qquad s, t \in T, \quad s < t$$

satisfies (S1). Because $\{A_t\}$ is stationary, (S2) also holds.

It remains to establish (S3). Since

$$g_{t-s} \equiv \mathrm{E}(x_{st}) = -\mathrm{E}\{\log (P_{st})_{11}\} \leqq -\mathrm{E}\left\{ \sum_{v=s+1}^{t} \log (A_v)_{11} \right\} = -(t-s)\mathrm{E}\{\log (A_1)_{11}\} < +\infty,$$

g_t exists. We show now that $\mathrm{E}\{\log \|A_1\|_1\} < \infty$, then that $g_t/t \geqq -\mathrm{E}\{\log \|A_1\|_1\} > -\infty$. Now

$$\mathrm{E}\{\log \|A_1\|_1\} = \mathrm{E}\left\{ \max_i \log \sum_j (A_1)_{ij} \right\}$$

$$\leqq \mathrm{E}\left\{ \max_{i,j} \log (d[A_1]_{ij}) \right\} = \log d + \mathrm{E}\left\{ \max_{i,j} \log [A_1]_{ij} \right\} < \infty.$$

Hence

$$-g_t = \mathrm{E}\{\log (P_{0t})_{11}\} \leqq \mathrm{E}\{\log \|P_{0t}\|_1\}$$

$$\leqq \sum_{v=1}^{t} \mathrm{E}\{\log \|A_v\|_1\} = t\mathrm{E}\{\log \|A_v\|_1\} < \infty$$

and

$$g_t/t \geqq -\mathrm{E}\{\log \|A_v\|_1\} > -\infty.$$

This proves (S3), so that Theorem B (Theorem 1 of Kingman (1973)) applies and gives the claimed result for $i = j = 1$.

As for the other elements of P_{st}, it is easy to see that, precisely as for ordinary matrix multiplication,

$$(P_{0t})_{ij} \geqq (A_1)_{i1}(P_{1,t-1})_{11}(A_t)_{1j}, \qquad t > 2,$$

whence, almost surely,

$$\liminf_{t\uparrow\infty} t^{-1} \log (P_{0t})_{ij} \geqq \liminf_{t\uparrow\infty} t^{-1} \log (A_1)_{i1} + \xi + \liminf_{t\uparrow\infty} t^{-1} \log (A_t)_{1j}.$$

Obviously the first term on the right of the inequality vanishes. We claim the last term vanishes also. To see this, observe that by stationarity (S2), $\mathrm{E}\{t^{-1} \log (A_t)_{1j}\} = \mathrm{E}\{t^{-1} \log (A_1)_{1j}\} \to 0$ as $t \uparrow \infty$, $t \in T$. Now for any $\varepsilon > 0$,

$$\sum_{t=1}^{\infty} P(-t^{-1} \log (A_t)_{1j} > \varepsilon) = \sum_{t=1}^{\infty} P(-\varepsilon^{-1} \log (A_1)_{1j} > t)$$

$$\leqq \mathrm{E}\{-\varepsilon^{-1} \log (A_1)_{1j}\} < \infty$$

since $\mathrm{E}\{\log (A_1)_{1j}\} > -\infty$. The identical argument also gives, for any $\varepsilon > 0$,

$$\sum_{t=1}^{\infty} P(+t^{-1} \log (A_t)_{1j} > \varepsilon) < \infty$$

since $\mathrm{E}\{\log (A_1)_{1j}\} < +\infty$. The Borel–Cantelli lemma (see, e.g., Breiman (1968, p. 41)) then implies

$$\lim_{t\uparrow\infty} t^{-1} \log (A_t)_{ij} = 0 \quad \text{with probability one.}$$

Thus

$$(*)\qquad \liminf_{t\uparrow\infty} t^{-1}\log(P_{0t})_{ij} \geqq \xi \quad \text{almost surely.}$$

Similarly,

$$(P_{0,t+2})_{11} \geqq (A_1)_{1i}(P_{1,t+1})_{ij}(A_{t+2})_{j1}, \qquad t\geqq 1$$

shows that

$$\limsup_{t\uparrow\infty} t^{-1}\log(P_{1,t+1})_{ij} \leqq \xi \qquad \text{almost surely.}$$

By stationarity,

$$\limsup_{t\uparrow\infty} t^{-1}\log(P_{0t})_{ij} \leqq \xi \quad \text{almost surely,}$$

which together with $(*)$ implies that the almost sure limit does not depend on i, j.

The proof of convergence in mean follows the same lines. □

Applied to the example of Ace Distributors, Theorem 2 implies that over a large number of days, the geometric mean maximum probability of successful shipment approaches a limiting value (under the assumed conditions), and this limiting value is the same for all pairs i, j of sites.

THEOREM 3. *Let $\{A_t\}_{t\in T}$ be a stationary sequence of random matrices from $\mathbb{R}_{++}^{d\times d}$. If $-\infty < \mathrm{E}\{\log(A_t)_{ij}\} < \infty$ for all $t\in T$, $1\leqq i, j\leqq d$, then*

$$\lim_{t\uparrow\infty} t^{-1}\log(P_{0t}^{\llcorner\cdot\times})_{ij} = \xi$$

exists with probability one, is independent of i and j, and has a finite mean

$$\mathrm{E}(\xi) = \lim_{t\uparrow\infty} t^{-1}\mathrm{E}\{\log(P_{0t}^{\llcorner\cdot\times})_{11}\} \equiv \gamma_{\llcorner\cdot\times}\{A_t\}.$$

If $\{A_t\}$ are i.i.d., then $\xi = \gamma_{\llcorner\cdot\times}\{A_t\}$ almost surely.

Proof. Let

$$x_{st} = \log(P_{st}^{\llcorner\cdot\times})_{11}.$$

Then

$$x_{su} \leqq x_{st} + x_{tu}, \qquad s<t<u, \quad s, t, u\in T,$$

$$g_t \equiv \mathrm{E}(x_{0t}) \leqq \sum_{v=1}^{t} \mathrm{E}\{\log(A_v)_{11}\} = t\mathrm{E}\{\log(A_1)_{11}\} < \infty,$$

and

$$g_t = \mathrm{E}(x_{0t}) \geqq \mathrm{E}\left\{\log\prod_{v=1}^{t}\min_{i,j}(A_v)_{ij}\right\}$$

$$= \sum_{v=1}^{t}\mathrm{E}\left\{\min_{i,j}\log(A_v)_{ij}\right\} = t\mathrm{E}\left\{\min_{i,j}\log(A_1)_{ij}\right\};$$

hence

$$g_t/t \geqq \mathrm{E}\left\{\min_{i,j}\log(A_1)_{ij}\right\} > -\infty.$$

These observations and the assumed stationarity of $\{A_t\}$ establish (S1)–(S3) and hence the applicability of Theorem B (the subadditive ergodic theorem) to x_{st}.

For the other elements of $P_{0t}^{\llcorner\cdot\times}$, use the inequalities

$$(P_{0t}^{\llcorner\cdot\times})_{ij} \leqq (A_1)_{i1}(P_{1,t-1}^{\llcorner\cdot\times})_{11}(A_t)_{1j},$$

$$(P_{0,t+2}^{\llcorner\cdot\times})_{11} \leqq (A_1)_{1i}(P_{1,t+1}^{\llcorner\cdot\times})_{ij}(A_{t+2})_{j1},$$

and argue as in the proof of Theorem 2. □

THEOREM 4. *Let $\{A_t\}_{t\in T}$ be a stationary sequence of random matrices from $\mathbb{R}^{d\times d}$. If $-\infty < \mathrm{E}\{(A_t)_{ij}\} < \infty$ for all $t \in T$, $1 \leqq i, j \leqq d$, then*

$$\lim_{t\uparrow\infty} t^{-1}(P_{0t}^{\ulcorner\cdot+})_{ij} = \xi$$

exists with probability one, is independent of i and j, and has a finite mean

$$\mathrm{E}(\xi) = \lim_{t\uparrow\infty} t^{-1}\mathrm{E}\{(P_{0t}^{\ulcorner\cdot+})_{11}\} \equiv \gamma_{\ulcorner\cdot+}\{A_t\}.$$

If $\{A_t\}$ are i.i.d., then $\xi = \gamma_{\ulcorner\cdot+}\{A_t\}$ almost surely.

Proof. $(\ulcorner\cdot+)$-multiplication is associative in $\mathbb{R}^{d\times d}$. For $s<t<u$, $s, t, u \in T$,

$$(P_{su}^{\ulcorner\cdot+})_{11} \geqq (P_{st}^{\ulcorner\cdot+})_{11} + (P_{tu}^{\ulcorner\cdot+})_{11}.$$

Then

$$x_{st} = -(P_{st}^{\ulcorner\cdot+})_{11}, \quad s,t\in T, \quad s<t$$

satisfies (S1). Because $\{A_t\}$ is stationary, (S2) also holds. Moreover,

$$g_t \equiv \mathrm{E}\{x_{0t}\} \leqq \sum_{v=1}^{t} \mathrm{E}\{-(A_v)_{11}\} = (-t)\mathrm{E}\{(A_1)_{11}\} < \infty,$$

so g_t exists, and

$$-g_t = \mathrm{E}\{(P_{0t}^{\ulcorner\cdot+})_{11}\} \leqq \mathrm{E}\left\{\sum_{v=1}^{t} \max_{i,j}(A_v)_{ij}\right\} = t\mathrm{E}\left\{\max_{i,j}(A_1)_{ij}\right\}$$

so

$$g_t/t \geqq -\mathrm{E}\left\{\max_{i,j}(A_1)_{ij}\right\} > -\infty.$$

This proves (S1)–(S3) and the claimed result for $i=j=1$ using Theorem B.

For the other matrix elements, use the inequalities

$$(P_{0t}^{\ulcorner\cdot+})_{ij} \geqq (A_1)_{i1} + (P_{1,t-1}^{\ulcorner\cdot+}\ulcorner\cdot+A_t)_{1j} \geqq (A_1)_{i1} + (P_{1,t-1}^{\ulcorner\cdot+})_{11} + (A_t)_{1j},$$

$$(P_{0,t+2}^{\ulcorner\cdot+})_{11} \geqq (A_1)_{1i} + (P_{1,t+1}^{\ulcorner\cdot+})_{ij} + (A_{t+2})_{j1},$$

and argue as in the proof of Theorem 2. □

Applied to the example of Best Manufacturing, the assumptions of Theorem 4 require that the operation of each of the d machines be involved in the cycle of each machine, since no element of A_t, $t \in T$, may be fixed at $-\infty$. Indeed, it is intuitively clear that two sets of noninteracting machines (represented by operating-time matrices A_t of block-diagonal form, with elements equal to $-\infty$ outside of the block-diagonal submatrices) could cycle at different rates. Theorem 4 says that, in a set of d fully interacting machines, the average duration of a cycle will approach a limit and this limit will be the same for all machines.

THEOREM 5. *Let $\{A_t\}_{t\in T}$ be a stationary sequence of random matrices from $\mathbb{R}^{d\times d}$. If $-\infty < \mathrm{E}\{(A_t)_{ij}\} < \infty$ for all $t \in T$, $1 \leqq i, j \leqq d$, then*

$$\lim_{t\uparrow\infty} t^{-1}(P_{0t}^{\llcorner\cdot+})_{ij} = \xi$$

exists with probability one, is independent of i and j, and has a finite mean

$$\mathrm{E}(\xi) = \lim_{t\uparrow\infty} t^{-1}\, \mathrm{E}\{(P_{0t}^{\llcorner\cdot+})_{11}\} \equiv \gamma_{\llcorner\cdot+}\{A_t\}.$$

If $\{A_t\}$ are i.i.d., then $\xi = \gamma_{\llcorner\cdot+}\{A_t\}$ almost surely.

Proof. ($\llcorner\cdot+$)-multiplication is associative in $\mathbb{R}^{d\times d}$. For $s<t<u$, $s, t, u \in T$,

$$(P_{su}^{\llcorner\cdot+})_{11} \leqq (P_{st}^{\llcorner\cdot+})_{11} + (P_{tu}^{\llcorner\cdot+})_{11}.$$

Then

$$x_{st} = (P_{st}^{\llcorner\cdot+})_{11}, \quad s, t \in T, \quad s<t$$

satisfies (S1) and (S2). Moreover $g_t \equiv \mathrm{E}\{x_{0t}\} \leqq t\mathrm{E}\{(A_1)_{11}\} < \infty$ exists and

$$g_t \geqq \sum_{v=1}^{t} \mathrm{E}\left\{\min_{i,j} (A_v)_{ij}\right\} = t\mathrm{E}\left\{\min_{i,j} (A_1)_{ij}\right\}$$

so

$$g_t/t \geqq \mathrm{E}\left\{\min_{i,j} (A_1)_{ij}\right\} > -\infty.$$

This proves (S1)–(S3) and, with the help of Theorem B, the claimed result for $i = j = 1$.

For the other matrix elements, use the inequalities

$$(P_{0t}^{\llcorner\cdot+})_{ij} \leqq (A_1)_{i1} + (P_{1,t-1}^{\llcorner\cdot+})_{11} + (A_t)_{1j},$$

$$(P_{0,t+2}^{\llcorner\cdot+})_{11} \leqq (A_1)_{1i} + (P_{1,t+1}^{\llcorner\cdot+})_{ij} + (A_{t+2})_{j1},$$

and argue as in the proof of Theorem 2. □

Applied to the transport costs of Ace Distributors, Theorem 5 says that the average (per stage) cost of a minimal-cost t-stage trip from site i to site j approaches a limit as t gets large, and this limiting minimal cost per stage is independent of the pair (i, j) of sites.

COROLLARY 6. *Under the hypotheses of Theorems 2 and 3, $\gamma_{\llcorner\cdot\times}\{A_t\} \leqq \gamma_{\ulcorner\cdot\times}\{A_t\}$. Under the hypotheses of Theorems 4 and 5, $\gamma_{\llcorner\cdot+}\{A_t\} \leqq \gamma_{\ulcorner\cdot+}\{A_t\}$. Under the hypotheses of Theorems 2–5, $\exp[\gamma_{\llcorner\cdot\times}\{A_t\}] \leqq \gamma_{\llcorner\cdot+}\{A_t\}$ and $\exp[\gamma_{\ulcorner\cdot\times}\{A_t\}] \leqq \gamma_{\ulcorner\cdot+}\{A_t\}$.*

Proof. For every sample path of $\{A_t\}$, $P_{0t}^{\llcorner\cdot g} \leqq P_{0t}^{\ulcorner\cdot g}$, where g takes the value $\times$ or $+$; hence the first two inequalities. Similarly, for every sample path of $\{A_t\}$ (on $\mathbb{R}_{++}^{d\times d}$), $(P_{0t}^{f\cdot\times})_{ij}^{1/t} \leqq t^{-1}(P_{0t}^{f\cdot+})_{ij}$, where f takes the value $\llcorner$ or $\ulcorner$, by the inequality of geometric and arithmetic means; hence the second two inequalities. □

The last remaining generalized multiplication among the examples in §4 is ($\ulcorner\cdot\llcorner$)-multiplication. Subadditive inequalities for individual matrix elements like those that begin the proofs of Theorems 4 and 5 are not possible for ($\ulcorner\cdot\llcorner$)-multiplication. For example, if

$$H = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

then $(H \ulcorner\cdot\llcorner H)_{11} = 1 > H_{11} + H_{11} = 0$ but $(I \ulcorner\cdot\llcorner I)_{11} = 1 < I_{11} + I_{11} = 2$.

Define, for $A \in \mathbb{R}^{d\times d}$,

$$\ulcorner(A)=\max_{i,j} a_{ij}, \qquad \sum(A)=\sum_{i,j} a_{ij}.$$

Then

$$\ulcorner(A\ulcorner\cdot\llcorner B)\leqq[\ulcorner(A)]\llcorner[\ulcorner(B)].$$

For

$$(A\ulcorner\cdot\llcorner B)_{ik}=(a_{i1}\llcorner b_{1k})\ulcorner(a_{i2}\llcorner b_{2k})\ulcorner\cdots\ulcorner(a_{id}\llcorner b_{dk})$$
$$\leqq a_{i1}\ulcorner a_{i2}\ulcorner\cdots\ulcorner a_{id}=\max_j a_{ij}\leqq\ulcorner(A)$$

and similarly

$$(A\ulcorner\cdot\llcorner B)_{ik}\leqq\max_j b_{jk}\leqq\ulcorner(B);$$

hence the claim $\ulcorner(A\ulcorner\cdot\llcorner B)\leqq[\ulcorner(A)]\llcorner[\ulcorner(B)]$. Consequently, for any sequence $\{A_t\}_{t\in T}$ of matrices over $\mathbb{R}^{d\times d}$,

$$\ulcorner(A_1\ulcorner\cdot\llcorner A_2\ulcorner\cdot\llcorner\cdots\ulcorner\cdot\llcorner A_t)\leqq\ulcorner(A_1).$$

If it is assumed that $-\infty<\mathrm{E}\{(A_1)_{ij}\}<\infty$, for $1\leqq i, j\leqq d$, then (by an argument like that in the second half of the proof of Theorem 2), $\lim_{t\uparrow\infty} t^{-1}\ulcorner(A_1)=0$ with probability one. Thus

$$\lim_{t\uparrow\infty} t^{-1}\ulcorner(P_{0t}^{\ulcorner\cdot\llcorner})\leqq 0 \quad \text{almost surely.}$$

If $A_t\in\mathbb{R}_+^{d\times d}$ for all t, with probability one, then $\ulcorner(P_{0t}^{\ulcorner\cdot\llcorner})\geqq 0$ almost surely, and hence

$$\lim_{t\uparrow\infty} t^{-1}\ulcorner(P_{0t}^{\ulcorner\cdot\llcorner})=0 \quad \text{almost surely.}$$

Now

$$\ulcorner(A)\leqq\sum(A)\leqq d^2\ulcorner(A) \quad \text{for } A\in\mathbb{R}_+^{d\times d}.$$

Hence, if $\{A_t\}_{t\in T}\in\mathbb{R}_+^{d\times d}$ almost surely, then also

$$\lim_{t\uparrow\infty} t^{-1}\sum(P_{0t}^{\ulcorner\cdot\llcorner})=0 \quad \text{almost surely.}$$

Applied to the example of Ace Distributors, the immediately preceding results mean that the maximum capacity of a t-stage trip is bounded above by the maximum capacity of a one-stage trip. Hence, for nonnegative capacities, the average capacity per stage (which is not an especially meaningful descriptive statistic) approaches zero with probability one, for every pair of sites.

In summary, the generalized products $P_{0t}^{f\cdot g}$ of stationary random matrices $\{A_t\}_{t\in T}$ display different asymptotic behavior as $t\uparrow\infty$, depending on the generalized matrix multiplication $f\cdot g$ and the space $F^{d\times d}$ from which the matrices come. According to Theorems 2 and 3, $\ulcorner\cdot\times$-products and $\llcorner\cdot\times$-products $P_{0t}^{f\cdot g}$ over $\mathbb{R}_{++}^{d\times d}$ asymptotically behave exponentially in t. According to Theorems 4 and 5, $\ulcorner\cdot+$-products and $\llcorner\cdot+$-products $P_{0t}^{f\cdot g}$ over $\mathbb{R}^{d\times d}$ asymptotically behave linearly in t. Finally, $\ulcorner\cdot\llcorner$-products over $\mathbb{R}_+^{d\times d}$ are absolutely bounded above and below and therefore vanish almost surely compared to t.

6. Generalized spectral radius of a fixed matrix. In the case of ordinary matrix multiplication, Theorem D implies, as already noted, that for every matrix $A\in\mathbb{R}_{++}^{d\times d}$

there is a spectral radius $\rho(A) \in \mathbb{R}_{++}$ such that, for all $1 \leqq i, j \leqq d$,

$$\lim_{t\uparrow\infty} [(A^t)_{ij}]^{1/t} = \rho(A).$$

Theorems 2–5 imply precisely analogous results in the special cases when the random matrices $\{A_t\}$ are all equal to a single fixed matrix A with probability one. However, unlike the spectral radius for ordinary $+\cdot\times$-multiplication, which can only be calculated exactly in a few special cases of A, the generalized spectral radii $\rho_{f\cdot g}(A)$ for those $f\cdot g$-multiplications covered by Theorems 2–5 can be calculated exactly by simple general formulas.

Define a *cycle* to be any cyclic permutation of any nonempty subset of $\{1, 2, \cdots, d\}$. For example, if $d=3$, then all the cycles are (1), (2), (3), (1, 2), (1, 3), (2, 3), (1, 2, 3) and (1, 3, 2). The cycle (2, 3, 1), which does not appear in this list, is equivalent to and represented by the cycle (1, 2, 3) in the list. Define the length $L=L(C)$ of cycle C to be the number of elements in the subset of $\{1, \cdots, d\}$ permuted by C. Thus $L(1, 2)=2$. Let S_d be the set of all possible cycles of any length L, $1 \leqq L \leqq d$, from $\{1, \cdots, d\}$.

If $C=(i_1, i_2, \cdots, i_L)$, where $1 \leqq L \leqq d$, define a *sum over C in A* $\sum_C A$ by

$$\sum_C A = a_{i_1 i_2} + a_{i_2 i_3} + \cdots + a_{i_{L-1} i_L} + a_{i_L i_1}.$$

When $L=1$ and $C=(i)$, $\sum_{(i)} A = a_{ii}$. Similarly, define a *product over C in A* $\prod_C A$ by

$$\prod_C A = \prod_{j=1}^{L} a_{i_j i_{j+1}}, \qquad i_{L+1} \equiv i_1.$$

Denote the t-fold $f\cdot g$-product of the matrix A with itself by $A^t_{f\cdot g}$. Thus $A^t_{+\cdot\times}$ is the ordinary tth power of A.

THEOREM 7. *Let $A \in \mathbb{R}^{d\times d}_{++}$. Then, for $1 \leqq i, j \leqq d$,*

$$\lim_{t\uparrow\infty} [(A^t_{\ulcorner\cdot\times})_{ij}]^{1/t} = \max\left\{\left(\prod_C A\right)^{1/L(C)} \,\middle|\, C \in S_d\right\},$$

$$\lim_{t\uparrow\infty} [(A^t_{\llcorner\cdot\times})_{ij}]^{1/t} = \min\left\{\left(\prod_C A\right)^{1/L(C)} \,\middle|\, C \in S_d\right\}.$$

Let $A \in \mathbb{R}^{d\times d}$. Then, for $1 \leqq i, j \leqq d$,

$$\lim_{t\uparrow\infty} t^{-1}(A^t_{\ulcorner\cdot+})_{ij} = \max\left\{\left(\sum_C A\right)\Big/ L(C) \,\middle|\, C \in S_d\right\},$$

$$\lim_{t\uparrow\infty} t^{-1}(A^t_{\llcorner\cdot+})_{ij} = \min\left\{\left(\sum_C A\right)\Big/ L(C) \,\middle|\, C \in S_d\right\}.$$

For example, let $d=3$ and

$$A = \begin{bmatrix} 54 & -77 & 13 \\ -56 & -19 & -46 \\ -31 & 34 & 49 \end{bmatrix}.$$

The possible cycles C and averaged sums over cycles in A are given in Table 1.
Thus $\lim_{t\uparrow\infty} t^{-1}(A^t_{\ulcorner\cdot+})_{ij} = 54$ and $\lim_{t\uparrow\infty} t^{-1}(A^t_{\llcorner\cdot+})_{ij} = -66.5$.

Proof of Theorem 7. Denote the four limits on the left of the equations in Theorem 7 by $\rho_{\ulcorner\cdot\times}(A)$, $\rho_{\llcorner\cdot\times}(A)$, $\rho_{\ulcorner\cdot+}(A)$ and $\rho_{\llcorner\cdot+}(A)$, respectively. We will first establish

TABLE 1

Cycle C	Averaged sum $(\sum_C A)/L(C)$
(1)	54
(2)	−19
(3)	49
(1, 2)	(−77 − 56)/2 = −66.5
(1, 3)	(13 − 31)/2 = −9
(2, 3)	(34 − 46)/2 = −6
(1, 2, 3)	(−77 − 46 − 31)/3 = −51.33 . . .
(1, 3, 2)	(13 + 34 − 56)/3 = −3

the claimed formula for the last of these, $\rho_{\llcorner\cdot+}(A)$. (To avoid confusion, note that when $A_t = A$ almost surely for all $t \in T$, $\gamma_{f\cdot\times}\{A_t\} = \log \rho_{f\cdot\times}(A)$, $f = \ulcorner, \llcorner, +$; but $\gamma_{f\cdot+}\{A_t\} = \rho_{f\cdot+}(A), f = \ulcorner, \llcorner$.)

Fix any i and j, $1 \leqq i, j \leqq d$, and pick any t that is large compared to d. Then there exists at least one sequence $i = j_0, j_1, j_2, \cdots, j_t = j$ such that $1 \leqq j_h \leqq d$, $h = 0, \cdots, t$ and

$$(A^t_{\llcorner\cdot+})_{ij} = \sum_{h=1}^{t} a_{j_{h-1}j_h}.$$

Call such a sequence $\{j_h\}$ a minimal sequence and the sum on the right in the last equation a minimal sum. Each minimal sequence may be partitioned into an initial, a middle, and a final sequence as follows. The initial sequence contains $i=j_0$ and all succeeding elements of the minimal sequence up to but not including the first repetition of a previous element of the minimal sequence. The final sequence contains $j=j_t$ and all preceding elements of the minimal sequence back to but not including the first repetition of an element that occurs later in the minimal sequence. The middle sequence contains all remaining elements of the minimal sequence. The initial and final sequences each contain at most d elements. Thus

$$t^{-1}(A^t_{\llcorner\cdot+})_{ij} = t^{-1} \sum_{\text{initial}} + t^{-1} \sum_{\text{middle}} + t^{-1} \sum_{\text{final}}$$

and as $t \uparrow \infty$ the initial and final sums on the right approach zero.

The middle sequence may be written as a disjoint sequence of cycles, except possibly for a finite number of elements. Since each cycle has length at most d, the number of such cycles goes to ∞ as $t \uparrow \infty$. Suppose that, as $t \uparrow \infty$, the middle sequence contained, infinitely often, a cycle $C' \in S_d$ such that

$$\left(\sum_{C'} A\right) \Big/ L(C') > \min\left\{\left(\sum_{C} A\right) \Big/ L(C) \,\Big|\, C \in S_d\right\}.$$

By bringing together all the elements of the minimal sequence belonging to any repetition of the cycle C', then replacing them by elements belonging to infinitely many repetitions of any cycle C'' such that

$$(**) \qquad \left(\sum_{C''} A\right) \Big/ L(C'') = \min\left\{\left(\sum_{C} A\right) \Big/ L(C) \,\Big|\, C \in S_d\right\}$$

(plus at most a finite number of other elements), we could strictly lower $t^{-1} \sum_{\text{middle}}$ for large enough t. Hence the minimal sequence containing C' infinitely often was not really minimal, contrary to assumption. Therefore, in the limit as $t\uparrow\infty$, the only cycles that can occur infinitely often in the middle sequence of a minimal sequence are those like C'' that satisfy $(**)$. This establishes that $\rho_{\llcorner\cdot+}(A)=\min\{(\sum_C A)/L(C) \mid C\in S_d\}$.

The proofs of the other three formulas in Theorem 7 replace either minimization by maximization or addition by multiplication and arithmetic mean by geometric mean, but are otherwise identical. □

Clearly, $\rho_{f\cdot\times}(A)>0$, $A\in\mathbb{R}_{++}^{d\times d}$, but $\rho_{f\cdot+}(A)$ may have any sign, $A\in\mathbb{R}^{d\times d}$, for $f=\ulcorner$, $f=\llcorner$.

This proof explains the perhaps puzzling result that for very long trips (large t), the average (per stage) cost in a minimal-cost trip for Ace's trucks from any site i to any site j is independent of i and j. The reason i and j have no influence is that the truck goes from i to a minimal-cost cycle and stays on one or another minimal-cost cycle (with at most a finite number of excursions for coffee) until shortly before t, when the truck leaves a minimal-cost cycle and travels to j. For large enough t, the costs of the initial segment from i to some minimal-cost cycle and the final segment from some minimal-cost cycle to j contribute negligibly to the average cost per stage.

Cuninghame-Green (1979, pp. 200–201) showed that, for $A\in\mathbb{R}^{d\times d}$, there exists a unique $\lambda\in\mathbb{R}$ and at least one $x\in\mathbb{R}^{d\times 1}$ such that $A\ulcorner\cdot+x=\lambda+x$ (addition on the right being elementwise); his λ is identical to our $\rho_{\ulcorner\cdot+}(A)$ as given by Theorem 7. Shortly after Cuninghame-Green originally published this result, and apparently independently, Vorob'ev (1963) produced a similar result for the generalized products $\ulcorner\cdot\times$ and $\llcorner\cdot\times$. Thus the generalized spectral radius $\rho_{\ulcorner\cdot+}(A)$ is also a generalized eigenvalue λ; similar results hold for the other generalized spectral radii considered here.

7. Exact formulas for the limiting growth rate. Theorems 2–5 establish the existence of limiting growth rates $\gamma_{f\cdot g}\{A_t\}$, for $f\cdot g=\ulcorner\cdot\times$, $\llcorner\cdot\times$, $\ulcorner\cdot+$, and $\llcorner\cdot+$, respectively, under various assumptions about $\{A_t\}$. Here, for each of these generalized products $f\cdot g$, we give an example of $\{A_t\}$ for which $\gamma_{f\cdot g}\{A_t\}$ is easily computed exactly and has a simple formula. To prepare for these formulas, we introduce some notation and a known almost sure limit law from the theory of extreme values.

Let $N(\mu,\sigma^2)$ denote a normally distributed random variable with mean μ and variance σ^2. For any two random variables X and Y, let $X\sim Y$ mean that X and Y have the same distribution. Let $X_1, X_2,\cdots$ denote a sequence of i.i.d random variables with $X_n\sim N(0,1)$, $n=1,2,\cdots$, and let

$$Z_n=\max\{X_1,\cdots,X_n\},$$

$$Y_n=\min\{X_1,\cdots,X_n\}.$$

By the symmetry of the normal distribution, $-Y_n\sim Z_n$.

THEOREM E (e.g., Galambos (1978, p. 228)).

$$P\left\{\lim_{n\uparrow\infty} Z_n/(2\log n)^{1/2}=1\right\}=1.$$

THEOREM 8. *Let $\{A_t\}_{t\in T}$ be a sequence of i.i.d. matrices from $\mathbb{R}^{d\times d}$ in which the elements $(A_t)_{ij}$ of each matrix are also i.i.d. and $(A_t)_{ij}\sim N(\mu,\sigma^2)$. Then, for $1\leqq i,j\leqq d$, with probability one,*

$$\gamma_{\ulcorner\cdot+}\{A_t\}=\lim_{t\uparrow\infty} t^{-1}(P_{0t}^{\ulcorner\cdot+})_{ij}=\mu+\sigma(2\log d)^{1/2},$$

$$\gamma_{\llcorner\cdot+}\{A_t\}=\lim_{t\uparrow\infty} t^{-1}(P_{0t}^{\llcorner\cdot+})_{ij}=\mu-\sigma(2\log d)^{1/2}.$$

Proof. Because (⌈ · +)-multiplication is associative,

$$(P_{0t}^{\ulcorner\cdot+})_{ij}=\max\left\{\sum_{h=1}^{t} a_{j_{h-1}j_h}\;\middle|\; j_0=i, j_t=j, 1\leqq j_h\leqq d, h=1,\cdots,t-1\right\},$$

where $a_{j_{h-1}j_h}$ is an element of A_h. There are d^{t-1} sums in the set on the right, and $(P_{0t}^{\ulcorner\cdot+})_{ij}$ is the maximum of these. Since $(A_h)_{ij}\sim N(\mu,\sigma^2)$,

$$t^{-1}\sum_{h=1}^{t} a_{j_{h-1}j_h}\sim N(\mu,\sigma^2/t)$$

$$\sim\mu+(\sigma/t^{1/2})N(0,1);$$

hence

$$t^{-1}(P_{0t}^{\ulcorner\cdot+})_{ij}\sim\mu+(\sigma/t^{1/2})Z_{d^{t-1}}.$$

Let $n=n(t)=d^{t-1}$. Then $t=1+(\log n)/(\log d)$. Hence

$$t^{-1}(P_{0t}^{\ulcorner\cdot+})_{ij}\sim\mu+\sigma Z_n/[1+(\log n)/(\log d)]^{1/2}.$$

As $t\uparrow\infty$, $n\uparrow\infty$, and Theorem E gives the formula claimed.

The proof for (⌊ · +)-multiplication uses $-Y_n\sim Z_n$. □

Theorem 8 tells Best Manufacturing (if the durations of operation of each of d machines may be supposed to be independently and identically normally distributed with mean μ and variance σ^2) that over a long period the average interval between cycles of each machine is $\mu+\sigma(2\log d)^{1/2}$. Compared to a manufacturing process with no variability ($\sigma=0$) but the same average duration μ of machine operation, the average intercycle interval is increased by $\sigma(2\log d)^{1/2}$. This penalty rises very slowly with increasing d (for example, increasing d from 20 to 40 increases the penalty from 2.4σ to 2.7σ approximately) but rises in direct proportion to σ. Thus, in an effort to improve (i.e., reduce) the average intercycle interval, reducing σ by half will be much more beneficial than reducing d by half, and similarly for any other proportional reduction. This conclusion was not obvious prior to analysis.

Similarly, Theorem 8 tells Ace Distributors (if the cost of a trip is i.i.d. $N(\mu,\sigma^2)$) that over a long period the average minimal cost per stage is $\mu-\sigma(2\log d)^{1/2}$, a savings of $\sigma(2\log d)^{1/2}$ over the average minimal cost per stage in the absence of variability in costs ($\sigma=0$). Here the variability σ and not the number d of sites is the major factor in Ace's savings.

THEOREM 9. *Let $\{A_t\}_{t\in T}$ be a sequence of i.i.d. matrices from $\mathbb{R}_{++}^{d\times d}$ in which the elements $(A_t)_{ij}$ of each matrix are also i.i.d. and* $\log(A_t)_{ij}\sim N(\mu,\sigma^2)$, *i.e., the elements of $(A_t)_{ij}$ are lognormally distributed. Then, for $1\leqq i, j\leqq d$, with probability one,*

$$\gamma_{\ulcorner\cdot\times}\{A_t\}=\lim_{t\uparrow\infty} t^{-1}\log(P_{0t}^{\ulcorner\cdot\times})_{ij}=\mu+\sigma(2\log d)^{1/2},$$

$$\gamma_{\llcorner\cdot\times}\{A_t\}=\lim_{t\uparrow\infty} t^{-1}\log(P_{0t}^{\llcorner\cdot\times})_{ij}=\mu-\sigma(2\log d)^{1/2}.$$

Proof. Apply Theorem 8 to $\{B_t\}_{t\in T}$ where $(B_t)_{ij}=\log(A_t)_{ij}$, $t\in T$, $1\leqq i$, $j\leqq d$. □

8. Open problems. How far can we relax the assumptions of Theorems 8 and 9 while retaining the conclusions? Any assumptions about i.i.d. $(A_t)_{ij}$ will do provided that the distribution of $t^{-1}\sum_{h=1}^{t} a_{j_{h-1}j_h}$ converges to the distribution of $N(\mu,\sigma^2/t)$ fast enough to justify the conclusion of the extreme value theorem, Theorem E. It should be possible to weaken the assumption that $(A_t)_{ij}$ are i.i.d. The problem is to make the details precise.

For ordinary matrix multiplication, e.g., Theorems C and D, the limit ξ is known as the leading Lyapunov exponent. Just as a single $d \times d$ matrix has a spectrum of d eigenvalues (possibly repeated), so an ordinary product of stationary random matrices has a spectrum of d Lyapunov exponents (Oseledec (1968); the theorem is stated and proved in a more readable way in Cohen, Kesten and Newman (1986)). Moreover, there is an intimate and surprising connection between the spectrum of Lyapunov exponents of ordinary products of certain i.i.d. matrices $\{A_t\}$ and the eigenvalues of the single random matrix $(A_1^T A_1)^{1/2}$ (Newman (1986)). To what extent do these results generalize to the $f \cdot g$-products considered in Theorems 2–5, or to other $f \cdot g$-products?

Is there a taxonomy of $f \cdot g$-products that could make the discovery and proof of subadditive ergodic theorems such as Theorems C, D and 2–5 more efficient? It would be natural to begin with $f, g \in \mathbf{O}$ and then extend to other binary operations f, g.

Finally, from the applied point of view, what generalized matrix products besides those studied here are important in science and management? How can the tools and results developed illuminate additional applied problems?

Acknowledgments. I thank Richard Gabriel of the Centre of Mathematical Statistics, Bucharest, Romania, for making me aware of the work of Cuninghame-Green (1979), and I thank Raymond Cuninghame-Green for helpful correspondence concerning his work.

I also thank Mr. and Mrs. William T. Golden for their hospitality.

REFERENCES

R. Bellman (1954), *Limit theorems for non-commutative operations* I, Duke Math. J., 21, pp. 491–500.

P. Bougerol and J. Lacroix (1985), *Products of Random Matrices with Applications to Schrödinger Operators*, Birkhauser, Boston.

L. Breiman (1968), *Probability*, Addison-Wesley, Reading, MA.

M. L. Cetlin (1952), *Application of matrix calculus to the synthesis of relay contact networks*, Dokl. Akad. Nauk SSSR (N.S.), 86, 525–528.

J. E. Cohen and C. M. Newman (1984), *The stability of large random matrices and their products*, Ann. Probab., 12, pp. 283–310.

J. E. Cohen, H. Kesten and C. M. Newman, eds. (1986), *Random Matrices and Their Applications*, Contemporary Mathematics, vol. 50, American Mathematical Society, Providence, RI.

R. Cuninghame-Green (1979), *Minimax Algebra*, Lecture Notes in Economics and Mathematical Systems 166, Springer-Verlag, Berlin, New York, Heidelberg.

H. Furstenberg and H. Kesten (1960), *Products of random matrices*, Ann. of Math. Statist. 31, pp. 457–469.

H. Furstenberg (1963), *Noncommuting random products*, Trans. Amer. Math. Soc., 108, pp. 377–428.

J. Galambos (1978), *The Asymptotic Theory of Extreme Order Statistics*, John Wiley, New York.

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

P. L. Hammer and S. Rudeanu (1968), *Boolean Methods in Operations Research and Related Areas*, Springer-Verlag, New York.

International Business Machines Corp. (1983), *APL Language Systems*, GC26-3847-5, 6th ed. IBM, San Jose, CA.

K. E. Iverson (1962), *A Programming Language*, John Wiley, New York.

K. H. Kim (1982), *Boolean Matrix Theory and Applications*, Marcel Dekker, New York.

J. F. C. Kingman (1968), *The ergodic theory of subadditive stochastic processes*, J. Roy. Statist. Soc. Ser. B, 30, pp. 499–510.

——— (1973), *Subadditive ergodic theory*, Ann. Probab., 1, pp. 883–909.

——— (1976), *Subadditive processes*, in Ecole d'Eté de Probabilités de Saint-Flour, P.-L. Hennequin ed., Lecture Notes in Mathematics, 539, Springer-Verlag, Berlin, pp. 165–223.

P. LANCASTER AND M. TISMENETSKY (1985), *The Theory of Matrices with Applications*, 2nd ed, Academic Press, Orlando, FL.

A. G. LUNC (1950), *Application of matrix Boolean algebra to the analysis and synthesis of relay-contact schemes*, Dokl. Akad. Nauk SSSR (N.S.), 70, pp. 421–423.

——— (1952), *Algebraic methods of analysis and synthesis of relay contact networks*, Izv. Akad. Nauk SSSR, Ser. Mat., 16, pp. 405–436.

GR. C. MOISIL (1960), *Sur une représentation des graphes qui interviennent dans l'économie des transports*, Com. Acad. R. P. Romine, 10, pp. 647–652.

C. M. NEWMAN (1986), *Lyapunov exponents for some products of random matrices: exact expressions and asymptotic distributions*, in Cohen, Kesten, and Newman, eds., 1986.

V. I. OSELEDEC (1968), *A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems*, Trans. Moscow Math. Soc., 19, pp. 197–231; Trudy Moskov. Mat. Obshch, 19, pp. 179–210 (in Russian).

G. PÓLYA AND G. SZEGÖ (1976), *Problems and Theorems in Analysis, Volume* I: *Series, Integral Calculus, Theory of Functions, 2nd ed*, Springer-Verlag, New York, Heidelberg.

N. N. VOROB'EV (1963), *An extremal matrix algebra*, Dokl. Akad. Nauk SSSR, 152, pp. 24–27.

SIAM REVIEW
Vol. 35, No. 1, p. 124, March 1993

006

ERRATUM: SUBADDITIVITY, GENERALIZED PRODUCTS OF RANDOM MATRICES AND OPERATIONS RESEARCH*

JOEL E. COHEN†

Amir Dembo of Stanford University and Yuval Peres of Yale University have pointed out to me an error in the proof of Theorem 8 (p. 83) and of Theorem 9 (p. 84), which depends on Theorem 8 [1]. Specifically, in the first paragraph of page 84, I assumed that for different trajectories $j_0, j_1, \ldots, j_i$, the normal random variables $\sum_{h=1}^{t} a_{j_{h-1}j_h}$ are independent. In fact, only for disjoint trajectories are these sums independent. In unpublished calculations, Dembo has further established that the four formulas for γ given in Theorems 8 and 9 are incorrect at least for $d = 2$. He also showed that the ratio between each formula given for γ and the corresponding true value converges to 1 as $d \to \infty$. Thus the formulas given provide good estimates for large d. The correct formulas for γ for any or all finite d are still unknown.

REFERENCE

[1] JOEL E. COHEN, *Subadditivity, generalized products of random matrices and operations research*, SIAM Rev., 30 (1988), pp. 69–86.

* Received by the editors August 3, 1992; accepted for publication September 21, 1992.
† Laboratory of Populations, Rockefeller University, New York, New York 10021-6399.