Z. Wahrscheinlichkeitstheorie verw. Geb. 18, 17–33 (1971)
© by Springer-Verlag 1971

Markov Chains and Summability Methods

JOHN G. KEMENY and J. LAURIE SNELL*

1. Introduction

The purpose of this paper is to discuss connections between certain classes of denumerable Markov chains and summability methods.

Let $A=\{A_{ij}\}$, $i, j=0, 1, 2, \ldots$, be an infinite matrix; and let $s=\{s_0, s_1, \ldots\}$ be a sequence, which we represent as a column vector. If the vector $t=As$ is well defined and converges to a, then A is said to *sum* s to a. If A sums all convergent sequences we say A defines a *summability method*. (The method is also denoted by A.) If A preserves limits for convergent sequences we say that A defines a *regular summability method*. The following theorem of Silverman and Toeplitz is basic to summability theory:

Theorem 1.1. A matrix A defines a regular summability method if and only if

(1.1) There is a B such that $\sum_{j=0}^{\infty} |A_{ij}| \leqq B$ for all i,

(1.2) $\lim_{i\to\infty} A_{ij}=0$ for all j,

(1.3) $\sum_{j=0}^{\infty} A_{ij} \to 1$ as $i\to\infty$.

We shall call a matrix A *triangular* if $A_{ij}=0$ for $j>i$ and $A_{ii}\neq 0$. If a summability method is defined by a triangular matrix we shall call it a *triangular method*. Most of the classical summability methods are defined by non-negative triangular matrices.

Let P be the transition matrix of a discrete time denumerable state Markov chain with states the non-negative integers. Consider the following problem: For what vectors s does $\lim_{n\to\infty} P^n s$ exist? In particular if s is the characteristic function of an infinite set E we are asking if $\lim_{n\to\infty} Pr_a[X_n\in E]$ exists for all starting states a. For a given starting state a, let $R_{nj}^{(a)}=P_{aj}^{(n)}$. Then for each a we wish $R^{(a)}s$ to have a limit.

Let P be either a transient or a null recurrent chain. Then $\lim_{n\to\infty} P_{aj}^{(n)}=0$ so $R^{(a)}$ has columns which tend to 0. That is, (1.2) is satisfied. Since $R^{(a)}$ has row sums one and is non-negative, (1.1) and (1.3) are also satisfied. Hence for each a, $R^{(a)}$ defines a regular summability method. Our original problem is thus transformed into the problem of determining if a sequence s is summed by each of a family $R^{(a)}$ of summability methods.

* The preparation of this paper was aided by the National Science Foundation.

It is easy to give examples of Markov chains where there is very little relation between the summability methods $R^{(a)}$ for different starting states a. However, the situation is quite different if we restrict the class of Markov chains to a class studied by Kemeny [8] which he has called *slowly spreading chains*. These are chains which have as state space the non-negative integers and which can move one step to the right but not more. We *shall prove that for these* chains the methods $R^{(a)}$ are equivalent in the following sense: for any two states a and b, $R^{(a)}$ sums s to l if and only if $R^{(b)}$ sums s to l. This enables us to reduce our original problem to one summability method, say $R^{(0)}$ which is triangular. Furthermore, if a summability method is triangular and can be associated with a Markov chain in the above manner, it must be $R^{(0)}$ for a slowly spreading chain. Since triangular methods play a special role in summability theory this gives added reason to give special attention to the case of slowly spreading chains.

The possibility of applying probability theory to the study of summability methods has already been demonstrated by Rényi [12] and Schmetterer [13]. Our purpose is to show how generally summability theory may be associated with Markov chains. We will take our terminology and notation for Markov chains from [10].

2. Some Summability Concepts

In this section we shall summarize some results from summability theory which we shall need later. Basic references for this subject are the books of Hardy [5], Cooke [3] and Zeller [15].

If A and B are two summability methods we say that B *includes* A if, whenever A sums a sequence s to a limit l, B also sums s to l. Methods A and B are *equivalent* if each includes the other. If B includes A but is not equivalent to A we say that B is *stronger* than A.

A regular summability method which sums only convergent sequences is called a *trivial method*. For example, the identity matrix I defines a trivial method. There is no summability method which sums all bounded sequences.

We mention now some special facts about triangular methods. Any triangular matrix A has a unique two sided inverse A^{-1} which is again triangular. The most useful result for comparing triangular methods is the following:

Theorem 2.1. *Let A and B be two regular triangular methods. Then B includes A if and only if BA^{-1} is a regular summability method and they are equivalent if BA^{-1} is trivial.*

In particular, by this theorem a regular triangular method is trivial if and only if A^{-1} is a regular method. If A has the property that $A_{ii}\to 0$, then the diagonal entries of A^{-1}, which are $1/A_{ii}$, approach ∞, so that A^{-1} cannot have bounded absolute row sums and A is not trivial. The converse is not true. It is possible to have a non-trivial method with diagonal entries not tending to zero.

We shall have occasion to deal with special triangular matrices of the form $A_{ij}=a_{i-j}$ for $i\geqq j$, where $a_0, a_1, \ldots$ is a given sequence with $a_0\neq 0$. Then $B=A^{-1}$ is again a matrix of the same form with $B_{ij}=b_{i-j}$ for a sequence $b_0, b_1, \ldots$. The generating functions $A(x)=\sum_{n=0}^{\infty} a_n x^n$ and $B(x)=\sum_{n=0}^{\infty} b_n x^n$ for these sequences are related by $A(x)B(x)=1$.

3. Renewal Sequences

From here on we let R stand for the $R^{(0)}$ for a Markov chain. Then $R_{nj}=P_{0j}^{(n)}$ and, in particular, $R_{n0}=P_{00}^{(n)}$. For any Markov chain the sequence $\{u_n\}$ defined by $u_n=P_{00}^{(n)}$ is an example of a *renewal sequence*. That is the sequence $\{u_n\}$ is such that $u_0=1$, $u_n=\sum_{j=0}^{n-1} f_{n-j}u_j$ where $\{f_n\}$ is a sequence with $f_0=0, f_n\geqq 0$ and $\sum f_n\leqq 1$. If $u_n=P_{00}^{(n)}$ for a Markov chain, then f_n is the probability, starting in state 0, that the first return to 0 is at time n.

A sufficient condition due to Kaluza [6] for a bounded sequence $\{u_n\}$ with $u_0=1$ to be a renewal sequence is that it satisfy

$$u_n^2\leqq u_{n-1}u_{n+1}. \tag{3.1}$$

In particular if $\{u_n\}$ is a moment sequence $u_n=\int_0^1 x^n\,dF(x)$ it satisfies this condition.

For slowly spreading chains we let $p_n=P_{n-1,n}$ and we assume that $p_n>0$ for $n=1,2,3,\ldots$ so that all states may be reached from state 0. Let $\beta_0=1$ and $\beta_{n+1}=\beta_n p_{n+1}$. Then $\beta_n=P_{0n}^{(n)}=R_{nn}$. Thus R_{nn} must be monotone decreasing. We let $\beta_\infty=\lim_{n\to\infty}\beta_n$, the probability that the chain "marches straight out". For a recurrent chain this limit is always zero, hence $R_{nn}\to 0$, and thus no recurrent chain yields a trivial method.

4. Slowly Spreading Chains

We now prove that the methods $R^{(a)}$ which arise from slowly spreading chains are equivalent. Before proving this theorem, we shall need to discuss some properties of slowly spreading chains.

Let P be the transition matrix of a slowly spreading chain. Define R by $R_{nj}=P_{0j}^{(n)}$ and S the "shift matrix" by $S_{i,i+1}=1$ and $S_{ik}=0$ otherwise. Then $RP=SR$, which determines the rows of R recursively. R is triangular and has strictly positive diagonal entries. It has a unique two sided inverse Q which is again triangular and

$$P=QSR \tag{4.1}$$

and

$$P^n=QS^nR. \tag{4.2}$$

If P has row-sums 1, which we assume, so do R and Q.

Theorem 4.1. *A slowly spreading chain is determined by its summability method.*

Proof. We have shown above that the summability method of a slowly spreading chain is R. Then Q is the unique triangular inverse of R, and P is determined by (4.1).

Theorem 4.2. *A triangular summability method A comes from a Markov chain if and only if $A\geqq 0$, A has row-sums 1, and $A^{-1}SA\geqq 0$.*

Proof. A comes from a Markov chain if $A_{nj}=P_{aj}^{(n)}$ for some state a. Triangularity means that P must be slowly spreading and $a=0$. Then $A=R$. Hence, it is non-negative and has row-sums 1. By (4.1), $P=A^{-1}SA$, and hence, this must be non-negative. But if all these conditions are satisfies, $P=A^{-1}SA$ is a Markov transition matrix and A is its summability method.

For slowly spreading chains the problem of finding non-negative, normalized eigenvectors

$$Ph^{(x)}=xh^{(x)}; \quad h_0^{(x)}=1, \tag{4.3}$$

was solved in [8]. For each x there is a unique normalized eigenvector $h^{(x)}$ associated with x and it is given by

$$h_i^{(x)}=\sum_{j=0}^{i} Q_{ij}x^j. \tag{4.4}$$

There is an $x_0, 0<x_0\leqq 1$, such that $h^{(x)}\geqq 0$ for all $x\geqq x_0$, and only for such x. The vectors $h^{(x)}$ are strictly positive if $x>x_0$; for $x=x_0$ there may be 0 components, but only if the states do not communicate. Also, $h^{(1)}=1$.

We now give a probabilistic interpretation for $h^{(x)}$ for $x>1$. The eigenvector $h^{(x)}$ is a regular function of the substochastic matrix $P^*=(1/x)P$. That is $P^*h^{(x)}=h^{(x)}$. Thus we can apply a systems theorem. Start P^* in state 0, and stop it if it reaches state i. Then

$$1=h_0^{(x)}=H_{0i}^*h_i^{(x)}$$

where H_{0i}^* is the probability starting in 0 that i is reached. Thus

$$h_i^{(x)}=1/H_{0i}^*. \tag{4.5}$$

That is $h_i^{(x)}$ is the reciprocal of the probability that the P^* process reaches state i. (Note that this argument makes use of the fact that a slowly spreading chain cannot jump over a state.) If the original process was stochastic, as we have tacitly assumed, then (4.5) is still correct for $x=1$, both sides being equal to 1.

By introducing the first-hitting probabilities $F_{ij}^{(n)}$, we can rewrite (4.5) as

$$1/h_i^{(x)}=\sum_{n=i}^{\infty} F_{0i}^{*(n)}=\sum_{n=i}^{\infty} F_{0i}^{(n)}\cdot(1/x)^n. \tag{4.6}$$

Theorem 4.3. *For a slowly spreading chain the summability methods $R^{(a)}$ for different starting states are equivalent.*

Proof. Let $R^{(a)}$ be the method obtained by starting in state a. Then by (4.2)

$$R_{nj}^{(a)}=P_{aj}^{(n)}=\sum_{l=0}^{a} Q_{al}R_{n+l,j}=\sum_{l=0}^{a} Q_{al}P_{0j}^{(n+l)}. \tag{4.7}$$

Define $u_n=Q_{a,a-n}$ for $n\leqq a$ and $u_n=0$ otherwise. Let U be the triangular matrix with $U_{nj}=u_{n-j}$. Finally let $\bar{R}=UR$. Then $\bar{R}_{n+a,j}=R_{nj}^{(a)}$ and we have simply added a rows to $R^{(a)}$ to make it into a triangular method. Adding a finite number of rows to a summability matrix has no effect on the matrix as to the sequences it sums. Thus $\bar{R}$ is equivalent to $R^{(a)}$. To prove that $\bar{R}$ and R are equivalent, by Theorem 2.1, it will be sufficient to prove that $\bar{R}R^{-1}=U$ is a trivial summability method.

U is a triangular matrix which has, from row a on, $Q_{aa}, \ldots, Q_{a0}$ starting from the diagonal entry, and zeros to the left. Thus U is clearly regular.

Let $V = U^{-1}$. Then $V_{nj} = v_{n-j}$ for a sequence $v_0, v_1, \ldots$. If $U(s) = \sum_{n=0}^{\infty} u_n s^n$ and $V(s) = \sum_{n=0}^{\infty} v_n s^n$, $U(s)\, V(s) = 1$. We use this to find V.

Now $U(s) = \sum_{n=0}^{a} Q_{a,a-n} s^n = \sum_{n=0}^{a} Q_{a,n} s^{a-n} = s^a h_a^{(1/s)}$ by (4.4).

$$V(s) = \frac{1}{U(s)} = \frac{1}{s^a} \frac{1}{h_a^{(1/s)}} = \frac{1}{s^a} \sum_{m=a}^{\infty} F_{0a}^{(m)} s^m$$

by (4.6). That is

$$V(s) = \sum_{m=0}^{\infty} F_{0a}^{(m+a)} s^m$$

or $v_m = F_{0a}^{(m+a)}$. Hence $v_m \geqq 0$, $\sum_{m=0}^{\infty} v_m = 1$ and $\lim_{m \to \infty} v_m = 0$. Thus $V = U^{-1}$ is regular, U is trivial, and this completes the proof.

5. Euler Methods

Consider the one-parameter family of slowly spreading chains defined by $P_{i,i+1} = p$ for some p, $0 < p \leqq 1$, and $P_{ii} = q = 1 - p$. This process represents the number of successes in independent trials with probability p for success. The corresponding summability method $R^{(p)}$ is given by

$$R_{nj}^{(p)} = \binom{n}{j} p^j q^{n-j}. \tag{5.1}$$

These are precisely the Euler summability methods.

The methods satisfy the simple algebraic identity

$$R^{(p_1)} R^{(p_2)} = R^{(p_1 p_2)} \tag{5.2}$$

for all real values of the parameter p, and

$$R^{(1)} = I. \tag{5.3}$$

Thus $Q^{(p)} = R^{(1/p)}$, and all computations are particularly simple.

We note that (1.2) requires that $p > 0$ and (1.1) that $p \leqq 1$. And these conditions suffice for regularity. Thus the regular Euler methods are precisely those obtained from Markov chains.

To compare $R^{(p)}$ with $R^{(\bar{p})}$, Theorem 2.1 tells us to study $R^{(p)} Q^{(\bar{p})} = R^{(p/\bar{p})}$. This is regular as long as $p \leqq \bar{p}$; its inverse is regular if $\bar{p} \leqq p$. Thus the first method is stronger than the second if and only if $p < \bar{p}$, and they are equivalent only if $p = \bar{p}$. We obtain a continuum of methods that grow weaker as p increases, and they terminate in the trivial method $p = 1$.

We will now show that from any slowly spreading chain P we may obtain a continuum of methods of varying strength. Let

$$\bar{P} = pP + qI, \quad p + q = 1. \tag{5.4}$$

$\bar{P}$ *will be slowly spreading as long as* $p>0$ and $\bar{P}\geqq 0$. If $\varepsilon=\inf_i P_{ii}$, this requires that

$$(5.5)\qquad 0<p\leqq\frac{1}{1-\varepsilon}.$$

Then a simple computation shows that

$$(5.6)\qquad \bar{R}=R^{(p)}R,$$

where $R^{(p)}$ is given by (5.1). We thus obtain a continuum of methods (5.6) for the parameter range (5.5). The original method is given by $p=1$.

To compare two such methods, let $\bar{\bar{R}}$ correspond to parameter value $\bar{p}$. Then

$$\bar{R}\bar{\bar{Q}}=R^{(p)}R\cdot QQ^{(\bar{p})}=R^{(p/\bar{p})}.$$

This is regular and non-trivial as long as $p<\bar{p}$. Thus the methods are strictly monotone decreasing in strength as p increases. For example, any one p yielding an Euler method gives rise to all the Euler methods by means of (5.4).

6. Balanced Methods

Let $b_0, b_1, \dots$ be a sequence of non-negative numbers with $b_0=1$. Let $\sigma_n=b_0+\cdots+b_n$. Let $R_{nj}=\frac{b_j}{\sigma_n}$ for $j\leqq n$ and 0 otherwise. Then R defines a regular summability method if $\sum b_j=\infty$. We shall call them *balanced methods*. We let $\beta_n=R_{nn}=b_n/\sigma_n$; then $Q=R^{-1}$ has $Q_{ii}=\frac{1}{\beta_i}$ and $Q_{i-i,i}=1-\frac{1}{\beta_i}$, and all other entries 0. If we form QSR we obtain

$$P_{ij}=(QSR)_{ij}=\begin{cases}\dfrac{b_j}{b_i}(\beta_i-\beta_{i+1}) & \text{if } j\leqq i\\[2ex] \dfrac{\beta_{i+1}}{\beta_i} & \text{if } j=i+1\\[2ex] 0 & \text{if } j>i+1.\end{cases}$$

If P is to be a Markov chain, β_n must be monotone decreasing. Then QSR is non-negative; that is, the necessary condition $\{\beta_n\}$ decreasing, is also sufficient for balanced methods to come from Markov chains.

Let us look also at the sequence $\{u_n\}$ with $u_n=P_{00}^{(n)}=R_{n0}=1/\sigma_n$. The sufficient condition of Kaluza for $\{u_n\}$ to be a renewal sequence was $u_{n-1}u_{n+1}-u_n^2\geqq 0$. That is $\sigma_n^2-\sigma_{n-1}\sigma_{n+1}=\sigma_n\sigma_{n+1}(\beta_n-\beta_{n+1})\geqq 0$. Thus we see that Kaluza's sufficient condition is necessary for balanced methods to come from chains.

In [9] Kemeny called these processes *balanced chains* and studied them in detail. He showed that the process is transient if and only if

$$\sum_n\frac{1}{\sigma_n}<\infty.$$

If this sum is infinite, it is recurrent and the stationary measure is proportional to $b=\{b_n\}$. Then it is ergodic if $\sum b_n<\infty$ and null if $\sum b_n=\infty$. The method is regular if $\sum b_n=\infty$, i.e. $\sigma_n\to\infty$. We note that

$$1-\beta_n=\frac{\sigma_{n-1}}{\sigma_n}$$

so that

$$\sigma_n=\frac{1}{\prod_{k=1}^{n}(1-\beta_k)}$$

for $n\geqq 1$. Thus in terms of the β's, a balanced chain yields a regular summability method if and only if

$$\sum_{k=0}^{\infty}\beta_k=\infty.$$

We ask next when R is trivial. That is, when is Q a regular method? Clearly from the form of Q this is simply a question of whether $1/\beta_n$ is bounded. We know that this sequence is monotone increasing, thus R is regular if and only if $\sum_k \beta_k=\infty$ and trivial if and only if $\beta_n\not\to 0$. The latter condition has an interpretation. β_n is the probability of taking n consecutive steps to the right. Thus a trivial balanced method is obtained from a chain only if there is a positive probability of marching deterministically to infinity (i.e. every step is a step to the right). Such a chain would, of course, be transient. Thus for this class of processes only some rather special transient chains can be trivial.

As a simple example of a balanced method we can choose all the b_n's equal to 1. This leads to the ordinary Cesàro summability method. The transition matrix for the associated chain has the form

$$\begin{aligned} P_{ij}&=\frac{1}{(i+1)(i+2)} && \text{if } j\leqq i\\ &=(i+1)/(i+2) && \text{if } j=i+1. \end{aligned}$$

In this case the higher order probabilities also have a quite simple form

$$\begin{aligned} P_{ij}^{(n)}&=\frac{n}{(i+n)(i+n+1)} && \text{if } j<i+n\\ &=(i+1)/(i+n+1) && \text{if } j=i+n. \end{aligned}$$

The problem of comparing two balanced methods was considered by Hardy [5] and also by Garabedian and Randels [4]. Hardy [5] gave the following two sufficient conditions for R to include $\bar{R}$.

Either

(6.1) $$\frac{b_n}{\bar{b}_n}\geqq\frac{b_{n+1}}{\bar{b}_{n+1}}$$

or

(6.2a) $$\frac{b_n}{\bar{b}_n}\leqq\frac{b_{n+1}}{\bar{b}_{n+1}}$$

and

(6.2b) $$\frac{\beta_n}{\bar{\beta}_n} < N \quad \text{for all } n.$$

Recall that $\beta_n = \frac{b_n}{\sigma_n}$.

Since, by (6.1), (6.2a) implies that $\bar{R}$ includes R, conditions (6.2) actually imply that R and $\bar{R}$ are equivalent.

Let $r_n = \frac{\beta_n}{\bar{\beta}_n}$. Then Garabedian and Randels [4] proved the following theorem:

Theorem 6.1. *A necessary and sufficient condition for R to include $\bar{R}$ is that for some N*

(6.3) $$\frac{1}{\sigma_n} \sum_{j=0}^{n-1} |\sigma_j r_j - \sigma_{j+1} r_{j+1}| + r_n < N$$

for all n.

Note that, by (6.1), if b_n is monotone decreasing it includes Cesàro summability. For example, if $b_n = \frac{1}{n+1}$ we obtain the so-called *logarithmic means.*

$$t_n = \frac{1}{\sigma_n}\left(s_0 + \frac{s_1}{2} + \cdots + \frac{s_n}{n+1}\right)$$

$$\sim \frac{1}{\log n}\left(s_0 + \frac{s_1}{2} + \cdots + \frac{s_n}{n+1}\right).$$

These means are in fact stronger than Cesàro means.

Since from the point of view of balanced chains the β_n's are the natural objects, it is useful to have sufficient conditions in terms of the ratios r_n. To this end we note the following result which follows immediately from Theorem 6.1.

Theorem 6.2. *If r_n is bounded and $\sum_{j=0}^{\infty} |r_j - r_{j+1}| < \infty$ then R includes $\bar{R}$. If in addition $\frac{1}{r_n}$ is unbounded, R is stronger than $\bar{R}$.*

It follows from Theorem 6.2 that if r_n is a decreasing sequence, R includes $\bar{R}$, and is stronger if $r_n \to 0$.

Examples. Let $\beta_n = \frac{1}{a n^{\varepsilon} + 1}$, $a > 0$, $0 < \varepsilon \leqq 1$. Then $\beta_n \to 0$ and $\sum \beta_n = \infty$, so that we have non-trivial regular methods. Let us call this the (a, ε) method.

$$r_n = \frac{\bar{a} n^{\bar{\varepsilon}} + 1}{a n^{\varepsilon} + 1}, \text{ bounded if and only if } \varepsilon \geqq \bar{\varepsilon}$$

$$r_n - r_{n+1} \sim \frac{(\varepsilon - \bar{\varepsilon})\, a \bar{a} n^{\varepsilon + \bar{\varepsilon}} + \varepsilon a n^{\varepsilon} - \bar{\varepsilon} \bar{a} n^{\bar{\varepsilon}}}{a^2 n^{1+2\varepsilon}}.$$

Hence if $\varepsilon > \bar{\varepsilon}$, $r_n - r_{n+1} \sim k/n^{1+(\varepsilon - \bar{\varepsilon})}$, $1/r_n$ is unbounded and, by Theorem 6.2, (a, ε) is stronger than $(\bar{a}, \bar{\varepsilon})$. If $\varepsilon = \bar{\varepsilon}$, $r_n - r_{n+1} \sim k/n^{1+\varepsilon}$, $1/r_n$ is bounded, and hence, the two methods are equivalent.

7. Nörlund Methods

These methods are defined in a manner quite similar to the definition of balanced methods. For a given sequence of non-negative numbers $c_0, c_1, \ldots$ with $c_0=1$, let $\gamma_n=c_0+\cdots+c_n$ and define R by

$$R_{nj}=\frac{c_{n-j}}{\gamma_n} \qquad j\leqq n$$
$$=0 \qquad \text{otherwise}.$$

To ensure that R gives a regular method we need only require that $\frac{c_n}{\gamma_n}\to 0$ as $n\to\infty$. We denote a Nörlund method determined by $c=\{c_n\}$ by (N, c). Unlike the case of balanced chains the matrix $Q=R^{-1}$ is quite complicated and it is difficult in practice to determine if a Nörlund method may be associated with a slowly spreading chain.

The most interesting class of Nörlund methods which can be shown to come from Markov chains is the class (C, α), of Cesàro means, for $\alpha>0$. These are Nörlund means with $c=\{c_n\}$ given by

$$c_n=\binom{n+\alpha-1}{n}.$$

In this case

$$R^{\alpha}_{nj}=\frac{\binom{n-j+\alpha-1}{n-j}}{\binom{\alpha+n}{n}}.$$

The transition matrix for the slowly spreading chain which yields R^{α} is given by P^{α} with

$$P^{\alpha}_{ij}=\frac{\alpha}{\alpha+1}\,\frac{\binom{\alpha+j-1}{j}}{\binom{\alpha+i+1}{i}} \qquad j\leqq i$$
$$=\frac{i+1}{\alpha+i+1} \qquad \text{for } j=i+1.$$

That this matrix P^{α} would yield R^{α} was conjectured by our student Jacob Bergmann. The proof that this is the case requires that one verify that

$$R^{\alpha}P^{\alpha}=SR^{\alpha}.$$

The computation is straightforward using the following binomial identity:

(7.1)
$$\sum_{k=j}^{i}\frac{\binom{i-k+\alpha-1}{i-k}}{\binom{k+\alpha+1}{k}}=\frac{\alpha+1}{i+\alpha+1}\,\frac{\binom{i-j+\alpha}{i-j}}{\binom{j+\alpha}{j}}.$$

The methods (C, α) become stronger as α increases.

The methods (C, α) are known to be equivalent to the Hölder methods (H, α). The methods $(C, 1)$ and $(H, 1)$ are the same but $(H, 2)$ differs from $(C, 2)$ and provides a simple example of a classical method where $A^{-1}SA$ is not non-negative and so is not associated with a Markov chain.

A rather special but interesting sequence which yields a Nörlund method associated with a Markov chain is the sequence $\{c_n\}=\{1, 1, 2, 3, 5, 8, \ldots\}$ of Fibonacci numbers. Using special facts about Fibonacci numbers our student Frank Hanna verified that QSR is non-negative and found the transition matrix to be

$$P_{ij}=\frac{\gamma_i}{\gamma_{i+1}}c_{i-j+1}-\frac{\gamma_{i-1}}{\gamma_i}c_{i-j}-\frac{\gamma_{i-2}}{\gamma_{i-1}}c_{i-j-1} \qquad \text{for } j\leqq i$$

$$=\frac{\gamma_i}{\gamma_{i+1}} \qquad \text{for } j=i+1.$$

8. Renewal Methods

We next consider a class of slowly spreading chains called "the basic example" in [10], and also known as renewal chains. They are characterized by the fact that they either take a step to the right or move all the way back to 0.

Let p_i be $P_{i-1,i}$ and $P_{i-1,0}=q_i$.

As usual $\beta_0=1$ and $\beta_i=\prod_{k=1}^{i} p_k$. Let $\beta_\infty=\lim \beta_i$. The process is transient if $\beta_\infty>0$, null if $\beta_\infty=0$ and $\sum\beta_i=\infty$, and ergodic if $\sum\beta_i<\infty$. The matrix Q has the form

$$Q_{ii}=1/\beta_i,$$

$$Q_{ij}=-\beta_{i-j-1}\,q_{i-j}/\beta_i \qquad \text{if } j<i.$$

To express R we introduce u_n recursively by

$$u_0=1 \quad \text{and} \quad u_{n+1}=\sum_{k=0}^{n}\beta_{n-k}\,q_{n+1-k}\,u_k$$

and

$$R_{nj}=u_{n-j}\,\beta_j.$$

The sequence $\{u_n\}$ defined by $u_n=R_{n0}$ is a renewal sequence and the sequence $\{f_n\}$ associated with $\{u_n\}$ is given by $f_n=\beta_{n-1}-\beta_n$. The condition that $f_n\geqq 0$ corresponds to the condition that the β_n's are decreasing. The condition $\sum f_n\leqq 1$ corresponds to $\beta_\infty\geqq 0$. It is this class of examples that shows that any renewal sequence $\{u_n\}$ can be $\{P_{00}^{(n)}\}$ for a Markov chain. If $U(s)=\sum_n u_n s^n$ is the generating function for the sequence $\{u_n\}$ and $B(s)=\sum_n \beta_n s^n$ is the generating function for the sequence $\{\beta_n\}$ the fact that

$$\sum_{j=0}^{n} R_{nj}=\sum_j u_{n-j}\,\beta_j=1$$

yields the relation

$$(1-s)\,B(s)\,U(s)=1 \tag{8.1}$$

between the generating functions $B(s)$ and $U(s)$. We can choose the sequence $\{\beta_n\}$ to be any positive monotone decreasing sequence with $\beta_0 = 1$ and the relation (8.1) will determine the sequence $\{u_n\}$. Choosing $\{u_n\}$ and then determining $\{\beta_n\}$ is much more difficult. However, we still have the sufficient condition of Kazula $u_n^2 \leqq u_{n-1} u_{n+1}$.

A summability method derived from a Markov chain is regular if the chain is null or transient. Thus a renewal method is regular if $\sum \beta_n = \infty$. It will be trivial if Q is regular. We know that this can only happen if the chain is transient, hence, $\beta_n \nrightarrow 0$. And in this case, we easily verify that Q is regular. Thus the non-trivial regular renewal methods are precisely those given by null chains.

We can obtain a one parameter family of examples by choosing

$$B(s) = (1-s)^{-p}$$

and

$$U(s) = (1-s)^{-q}, \qquad \text{with } p+q=1.$$

Then, by the binomial theorem,

$$\beta_j = \binom{p+j-1}{j}$$

$$u_j = \binom{q+j-1}{j}.$$

The resulting summability method has

$$R_{nj} = \binom{q+n-j-1}{n-j}\binom{p+j-1}{j}.$$

It is interesting to compare renewal methods with balanced methods. Recall that $\beta_n \to 0$ and $\sum \beta_n = \infty$ was the necessary and sufficient condition for a balanced method to give a non-trivial regular method. But the same is true for renewal chains.

This suggests comparing the two methods. Consider a null renewal chain determined by $\{\beta_n\}$. Let R be its summability method. Let $\bar{R}$ be the method of the balanced chain having the same β_n's. Then, by Theorem 2.1, R includes $\bar{R}$ if $R\bar{Q}$ is regular. For a balanced chain $\bar{Q}_{ii} = \dfrac{1}{\beta_i}$ and $\bar{Q}_{i,i-1} = 1 - \dfrac{1}{\beta_i}$. Also $R_{nj} = u_{n-j}\beta_j$, and $u_n \to 0$ since the chain is null. Thus,

$$(R\bar{Q})_{nj} = u_{n-j} - u_{n-j-1}(1-\beta_{j+1}).$$

Condition (1.1) is the only condition for regularity that is not evident. For this we need

$$\sum_{j=0}^{n} |u_{n-j} - u_{n-j-1}(1-\beta_{j+1})| < N$$

for all n. It is sufficient for this to have

(8.2) $$\sum_{j=1}^{n} |u_j - u_{j-1}| < N \qquad \text{for all } n.$$

This will, of course, be the case if the u_n's are monotone decreasing as if often true, for example for moment sequences. Thus renewal methods very often include the corresponding balanced method. We do not know if (8.2) is true for all renewal sequences. The result that (8.2) is a sufficient condition for R to include $\bar{R}$, is due to our student M. Berg.

9. Generalized Nörlund Methods

These methods introduced by Borwein [1] are defined as follows: Let $b=\{b_0, b_1, \ldots\}$ and $c=\{c_0, c_1, \ldots\}$ be two sequences of non-negative numbers with $b_0=c_0=1$ and $\sum b_n=\sum c_n=\infty$. Let $\tau_n=b_0\, c_n+b_1\, c_{n-1}+\cdots+b_n\, c_0$. Define R by $R_{nj}=\frac{c_{n-j}\, b_j}{\tau_n}$ for $j\leqq n$, and 0 otherwise. If R is to be associated with a Markov chain we must have β_n decreasing. That is

$$\frac{b_n}{\tau_n}\geqq\frac{b_{n+1}}{\tau_{n+1}}. \tag{9.1}$$

If $c_n=1$ for all n, then we have the balanced methods, and we have seen that (9.1) is necessary and sufficient. If $b_n=1$ for all n, we have the Nörlund methods, and here this condition is not sufficient; but we do get an interesting class (C, α) which are associated with Markov chains. If $\tau_n=1$ then (9.1) is again both necessary and sufficient and we get the renewal method with $\beta_n=b_n$ and $u_n=c_n$.

Borwein introduced the generalized Nörlund methods with

$$b_j=\binom{j+\alpha-1}{j},$$

$$c_j=\binom{j+\beta}{j},$$

$$\tau_n=\binom{n+\alpha+\beta}{n}.$$

He calls these generalized Cesaro means and denotes them by (C, α, β). If $\alpha>0$ and $\beta>-1$ these methods are regular and Borwein proved that (C, α, β) is equivalent to (C, α). The renewal examples we considered were $(C, \alpha, -\alpha)$ with $\alpha=p$, hence $0<\alpha<1$. Thus these methods are equivalent to the corresponding (C, α) methods.

10. Boundary Theory

The Martin boundary points of a transient Markov chain correspond to the minimal, non-negative regular functions. Since a slowly spreading chain has only one such function, its boundary is trivial. However, the space-time process of such a chain is quite interesting. This is obtained by taking as states time-position pairs (n, j). Such a chain is transient even if the original chain was recurrent. Its boundary gives information about the asymptotic behavior of the original chain. The space-time boundary was solved for slowly spreading chains in [8] under the assumption

(10.1) $$P_{ii} \geqq \varepsilon > 0 \quad \text{for all } i.$$

Then the boundary is a continuum $x_0 \leqq x \leqq +\infty$ with $0 < x_0 \leqq 1$. Some condition is necessary to assure a boundary which is a continuum. For example, one can show that the boundary of balanced chains can be identified with the non-negative integers. Assume now that (10.1) is satisfied. The minimal space-time function corresponding to x is the function $h(n,j) = \frac{h_j^{(x)}}{x^n}$ where $h^{(x)}$ is given by (4.4). The function corresponding to $+\infty$ is the limit of these functions which is δ_{nj}/β_j. We shall find it more convenient to use as boundary parameter.

(10.2) $$p = \frac{x - x_0}{x} \quad 0 \leqq p \leqq 1; \qquad x = \frac{x_0}{1-p}.$$

For each p we can define a modified slowly spreading chain by

(10.3) $$P_{ij}^{(p)} = \frac{P_{ij} h_j^{(x)}}{x\, h_i^{(x)}}; \qquad R_{nj}^{(p)} = \frac{1}{x^n} R_{nj} h_j^{(x)}$$

where x and p are connected by (10.2). This is well defined and stochastic as long as $h^{(x)} > 0$ and $P h^x = x h^x$ holds. This will be the case if $x_0 < x < \infty$, and if $x = x_0$ if the states communicate. Thus (10.3) gives us a new slowly spreading chain for $0 < p < 1$ and possibly also for $p = 0$.

It is natural for several reasons to consider the class $P^{(p)}$ for $0 < p < 1$ obtained from a given P. First it follows from general Martin boundary theory that they all have the same space-time boundary. The parametrization chosen is also natural from the point of view of summability theory since it is independent of the "original" R. To see this, suppose, for example, we had chosen an $\bar{R}$ corresponding to $x = x_1 > x_0$. That is

$$\bar{P}_{ij} = \frac{P_{ij} h_j^{(x_1)}}{x_1\, h_i^{(x_1)}},$$

$$\bar{R}_{nj} = \frac{1}{x_1^n} R_{nj} h_j^{(x_1)}.$$

$\bar{P}$ has as eigen-functions

$$\bar{h}^{(x)} = \frac{h^{(x x_1)}}{h^{(x_1)}} \quad \text{for } x > \bar{x}_0 = \frac{x_0}{x_1}.$$

Consider now $\bar{R}^p$ for $0 < p < 1$.

$$\begin{aligned} \bar{R}_{nj}^p &= \frac{1}{x^n} \bar{R}_{nj} \bar{h}_j^{(x)} \\ &= \frac{1}{x^n} \left[\frac{1}{x_1^n} R_{nj} h_j^{(x_1)} \right] \frac{h_j^{(x x_1)}}{h_j^{(x_1)}} \\ &= \frac{1}{(x x_1)^n} R_{nj} h_j^{(x x_1)}. \end{aligned}$$

And $p=(x-\bar{x}_0)/x=(x x_1-x_0)/(x x_1)$, thus it corresponds to the parameter value $x x_1$ for the original process. Therefore, $\bar{R}^p=R^p$, and we have shown that the parametrization chosen for p makes R^p independent of the choice of R.

Let us consider the simple sums-of-independent-random variables example with $P_{ii}=q$, $P_{i,i+1}=p$. We saw in Section 5 that these chains yield the Euler methods. Let us start with $p=q=\frac{1}{2}$. Then $R_{nj}=\binom{n}{j}(\frac{1}{2})^n$. The eigen-functions are $h_i^{(x)}=(2x-1)^i$ for $x\geqq\frac{1}{2}$. Thus $x_0=\frac{1}{2}$. Introducing a parameter p by means of (10.2) yields $h_i^{(x)}=(p/q)^i$ and $R_{nj}^p=\binom{n}{j}p^j q^{n-j}$. Thus the parametrization of the Euler continuum of methods is an example of a continuum obtained from boundary theory. In this case, the methods become progressively weaker as p increases, $0<p\leqq 1$. In particular the class of chains has the desirable property that if any one of them assigns a limit to $P_{iE}^{(n)}$, the probability after n steps of being in a set E, as $n\to\infty$, then any other chain for which this limit exists assigns the same limit.

We would like to believe that it is generally true that in each boundary continuum of methods $R^{(p)}$, $R^{(p_1)}$ is stronger than $R^{(p_2)}$ if $p_1<p_2$. Unfortunately we have only partial results in this direction.

Theorem 10.1. *If* (10.1) *holds, then* $R^{(p)}$ *is non-trivial for* $p<1$.

Proof. Assume $p<1$. Consider $R_{nn}^{(p)}$. This is the probability that $R^{(p)}$ takes n steps to the right. Its limit is the probability that the space-time process marches out along the diagonal (n,n), and hence that it goes to the boundary point $x=\infty$. But a process conditioned to get to one boundary point cannot go to another. Hence $R_{nn}^{(p)}\to 0$. But this means that $Q_{nn}^{(p)}$ would have entries tending to infinity and could not be regular. Thus $R^{(p)}$ is not trivial.

When the boundary is not a continuum the methods $R^{(p)}$ may all be trivial. Consider, for example, the case of a balanced chain. As we mentioned earlier the space-time boundary here turns out to be a denumerable set. For a balanced chain Q has all entries 0 except $Q_{ii}=1/\beta_i$ and $Q_{i,i-1}=(1-\beta_i)/\beta_i$. For Q to be regular it is then necessary and sufficient for $1/\beta_i$ to be bounded or $\beta_\infty>0$. But

$$\beta_n^{(x)}=1-\frac{1-\beta_n}{x},$$

$$\beta_\infty^{(x)}=1-x_0/x.$$

Thus $\beta_\infty^{(x)}=0$ if and only if $x=x_0$. Thus the method is trivial for all $p>0$.

For our next result we need a key lemma.

Lemma 10.2. *If* $\beta_\infty=0$ *and* $x>1$, *then* $\lim_n \frac{x^n}{h_n^{(x)}}=0$.

Proof. Let $q_k=1-P_{k-1,k}$. Then $\beta_\infty=\prod_{k=1}^{\infty}(1-q_k)=0$, hence $\sum q_k=+\infty$. Let $P^*=(1/x)P$, as in Section 4.

$$\frac{x^n}{h_n^{(x)}}=x^n H_{0n}^*=\prod_{k=0}^{n-1}(x H_{k,k+1}^*).$$

Considering separately the probability of reaching $k+1$ in one step or in more steps, we obtain

$$H^*_{k,k+1} \leqq \frac{1}{x} p_{k+1} + \frac{1}{x^2} q_{k+1},$$

$$x H^*_{k,k+1} \leqq p_{k+1} + \frac{1}{x} q_{k+1} = 1 - q_{k+1}\left(1 - \frac{1}{x}\right).$$

Hence,

$$\lim_n \frac{x^n}{h_n^{(x)}} = \prod_{k=0}^{\infty} [x H^*_{k,k+1}] \leqq \prod_{k=0}^{\infty} \left[1 - q_{k+1}\left(1 - \frac{1}{x}\right)\right] = 0,$$

since $\sum q_{k+1}\left(1 - \frac{1}{x}\right)$ diverges. This completes the proof.

While we cannot show that $R^{(p_1)}$ is stronger than $R^{(p_2)}$ for $p_1 < p_2$, we have seen that this is true for $p_2 = 1$, and we will now show that $R^{(p_2)}$ cannot include $R^{(p_1)}$. Since the choice of R is arbitrary, we choose it as $R^{(p_1)}$ in the theorem.

Theorem 10.3. *If* $\beta_\infty = 0$, *and* $x > 1$, *then* $R^{(x)}$ *does not include* R.

Proof. Suppose that $R^{(x)}$ includes R. Then $R^{(x)}Q$ is regular and its diagonal entries must be bounded. But

$$[R^{(x)}Q]_{nn} = \frac{1}{x^n} R_{nn} h_n^{(x)} Q_{nn} = \frac{h_n^{(x)}}{x^n},$$

which is unbounded as a consequence of the Lemma.

11. Perfect Methods

It would seem desirable for summability methods to have the following property. If A is a fixed method, it should be the case that when a method B sums every sequence A does, then A and B should give the same limit to sequences they both sum. It is not the case that all regular summability methods have this property. We shall say that a method is *perfect* if it does have this property. A fundamental theorem of Brudno [2] states that if we restricted ourselves to bounded sequences all methods would be perfect. That is, Brudno proved that if a regular method B sums all the bounded sequences a regular method A sums, then for bounded sequences which they both sum they give the same limit.

For triangular methods, it follows from results of Mazur [11] that a triangular method A is perfect if and only if it satisfies the following condition:

Condition M: If α is a row vector such that $\sum |\alpha_j| < \infty$ and $\alpha A = 0$ then $\alpha = 0$.

We can give a more probabilistic version of condition M when dealing with an R from a Markov chain. Let the chain be started in state 0 and stopped at a random time τ independent of the process. Let α be the distribution of the stopping time. Then α is a probability measure and the probability that the process is stopped at state j is

$$\sum_n \alpha_n P_{0j}^{(n)} = \mu_j.$$

Condition M is equivalent to saying that if for two stopping times τ and τ' the final distributions μ and μ' are the same than the distribution α and α' of τ and τ' must be the same.

A sufficient condition for condition M is the following. Let U^+ be Q with the negative entries replaced by 0. Then condition M is satisfied when U^+ has bounded columns. This is the case for balanced methods, methods (C, α), and renewal methods. It is the case for Euler methods only when $p \geqq \frac{1}{2}$. The general case for the Euler method was proven by Mazur [11] using analytic methods. We shall now show that it is possible to have a method which comes from a Markov chain which is not perfect. We shall show that this can occur for a class of chains called *random walks* by Karlin and McGregor [7]. These are slowly spreading chains which can move at most one step to the left. Let $p_i > 0$, $r_i \geqq 0$, $q_{i+1} > 0$ be such that $q_i + r_i + p_i = 1$. Define $P_{i,i-1} = q_i$ for $i > 0$, $P_{ii} = r_i$ and $P_{i,i+1} = p_i$ for $i \geqq 0$. Define $\pi_0 = 1$ and for $n \geqq 1$,

$$\pi_n = \frac{p_0 p_1 \cdots p_{n-1}}{q_1 q_2 \cdots q_n}.$$

Karlin and McGregor proved the representation theorem

$$P_{ij}^{(n)} = \pi_j \int_{-1}^{1} x^n Q_i(x) Q_j(x) \, d\psi(x).$$

Where the polynomials $\{Q_n\}$ are of degree n and are orthogonal with respect to the measure $\psi(x)$ on the interval $[-1, 1]$, $\psi(x)$ is called the spectral measure.

To find an example of an R that is not perfect we must find a chain such that there is an α, not 0, with $\sum |\alpha_i| < \infty$, and

$$\sum_n \alpha_n P_{0j}^{(n)} = 0$$

for all j. That is

$$\sum_n \alpha_n \int_{-1}^{1} x^n Q_j(x) \, d\psi(x) = 0$$

for all n. If we let $F(x) = \sum \alpha_n x^n$ be the generating function for α we see that we must have

$$\int_{-1}^{1} F(x) Q_j(x) \, d\psi(x) = 0$$

for all j.

But this states that $F(x)$ is a function with Fourier coefficients with respect to $\{Q_n\}$ equal to 0. From the theory of orthogonal polynomials this means that F must be 0 almost everywhere with respect to ψ. Thus if ψ has an interval of positive measure F would have to be zero on an interval and this would imply $F(x) \equiv 0$, since it defines an analytic function in the unit circle. The usual examples of random walks have measures ψ which have absolutely continuous parts and hence the corresponding methods are perfect.

It is possible, however, to use the Karlin-McGregor representation theorem to show that not all methods that come from random walks are perfect. To see this we first observe that according to Karlin and McGregor, any symmetric measure ψ on $[-1, 1]$ is the spectral measure of a random walk. Choose a function

$F(z)$ such that $F(z)=\sum \alpha_n z^n$ with $\sum |\alpha_n| < \infty$ and $F(z)=0$ on a sequence of points $\{\pm b_n\}$ with 1 and -1 as limit points. For example

$$F(z)=\sin\frac{1}{(1-z^2)^{\varepsilon}}\exp\frac{-1}{\sqrt{z^2-\varepsilon}}$$

for ε sufficiently small has these properties.[1] Then if ψ is a measure with support the points $\{\pm b_n\}$ the corresponding random walk will not be perfect.

References

1. Borwein, D.: On products of sequences. J. London math. Soc. **33**, 352-357 (1958).
2. Brudno, A.: Summation of bounded sequences by matrices. Mat. Sbornik, n. Ser. **16**, 191-247 (1945).
3. Cooke, R.G.: Infinite matrices and sequence spaces. London: Macmillan 1950.
4. Garabedian, H.L., Randels, W.C.: Theorems on Riesz means. Duke math. J. **4**, 529-533 (1938).
5. Hardy, G.H.: Divergent series. London: Oxford Press 1949.
6. Kaluza, T.: Über die Koeffizienten reziproker Potenzreihen. Math. Z. **28**, 161-170 (1928).
7. Karlin, S., McGregor, J.: Random walks. Illinois J. Math. **3**, 66-81 (1959).
8. Kemeny, J.G.: Representation theory for denumerable Markov chains. Trans. Amer. math. Soc. **125**, 47-62 (1966).
9. – Slowly spreading chains of the first kind. J. math. Analysis Appl. **15**, 295-310 (1966).
10. – Snell, J.L., Knapp, A.W.: Denumerable Markov chains. Princeton, N.J.: Van Nostrand 1966.
11. Mazur, S.: Eine Anwendung der Theorie der Operationen bei der Untersuchung der Toeplitzschen Limitierungsverfahren. Studia math. **2**, 40-50 (1930).
12. Rényi, A.: Summation methods and probability methods. Publ. math. Inst. Hungar. Acad. Sci. **4**, 389-397 (1959).
13. Schmetterer, L.: Wahrscheinlichkeitstheoretische Bemerkungen zur Theorie der Reihen. Arch. der Math. **14**, 311-316 (1963).
14. Spitzer, F.L.: Principles of random walk. Princeton, N.J.: Van Nostrand 1964.
15. Zeller, K.: Theorie der Limitierungsverfahren. Berlin-Göttingen-Heidelberg: Springer 1958.

[1] This example was provided for us by Mrs. Dona Strauss.

Prof. J.G. Kemeny and
Prof. J.L. Snell
Dept. of Mathematics
Dartmouth College
Hanover, New Hampshire 03755
USA

(Received September 23, 1969)