(1)

LIMIT THEOREMS FOR NONCOMMUTATIVE PROCESSES — II:
ON A GENERALIZATION OF THE STIELTJES INTEGRAL

Richard Bellman

P-772

28 November 1955

The RAND Corporation
1700 MAIN ST. • SANTA MONICA • CALIFORNIA

11 pp

Summary

~~In~~ this paper presents [redacted] two generalizations of the Riemann-Stieltjes integral arising from the study of positive definite matrices. ( )

LIMIT THEOREMS FOR NON-COMMUTATIVE PROCESSES—II ON A GENERALIZATION OF THE STIELTJES INTEGRAL*

By

Richard Bellman

§1. Introduction

If $f(t)$ is a continuous function of t over the interval $[0, 1]$ and $g(t)$ is a bounded monotone increasing function of t over the same interval, we know that the sum

$$(1) \qquad S_N = \sum_{i=0}^{N-1} f(t_i) \left[g(t_{i+1}) - g(t_i) \right],$$

where $0 = t_0 < t_1 < t_2 < \cdots < t_N = 1$, converges to a linear functional, which we may write $\int_0^1 f(t)dg$, as $N \to \infty$ and $\max_i (t_{i+1} - t_i) \to 0$.

This integral, the Riemann-Stieltjes integral, has been generalized in many different directions, cf. Bochner, [3]. We propose here to discuss two new generalizations arising from the study of positive definite matrices.

§2. First Generalization

Let $x(t)$ be a matrix-function of t for $0 \leq t \leq 1$ possessing the property that $x(t_2) - x(t_1)$ is non-negative definite whenever $1 \geq t_2 > t_1 \geq 0$. Let us now consider successive sub-divisions

*The first paper of this series is [2].

of the interval $[0, 1]$ which are refinements of the preceding, and to simplify the notation—since the essential difficulties do not lie in this direction—assume that $t_{i+1} - t_i = 1/2^N$ for the N^{th} sub-division.

Define $\lambda_1(t_i), \lambda_2(t_i), \cdots, \lambda_N(t_i)$ to be the characteristic values of the matrix $x(t_{i+1}) - x(t_i)$ arranged in decreasing order of magnitude, $\lambda_1(t_i) \geq \lambda_2(t_i) \geq \cdots \geq \lambda_N(t_i) \geq 0$.

Our first result is

Theorem 1. *Let* $f(t)$ *be a continuous function of* t *in* $[0, 1]$. *For each* k *the sum*

(1) $$S_N = \sum_{i=0}^{N-1} f(t_i) \lambda_k(t_i)$$

approaches a linear functional, which we write $\int_0^1 f(t) d\lambda_k$, *as* $N \rightarrow \infty$.

§3. Proof

The first part of our proof consists of showing that it is sufficient to prove the theorem for the case where $f(t)$ is a constant.

Divide the interval $[0, 1]$ into the 2^k intervals $[r2^{-k}, (r+1)2^{-k}]$, $r = 0, \cdots, 2^k - 1$, where k is chosen sufficiently large so that

(1) $$|f(t) - f(t_i)| \leq \epsilon \text{ for } r2^{-k} \leq t, t_i \leq (r+1)2^{-k}.$$

Then, for any $N > 2^k$,

$$(2)\qquad \left| S_N - f(0)\sum_{l=0}^{k_1-1}\lambda_k(t_l) - f(2^{-k})\sum_{l=k_1}^{2k_1-1}\lambda_k(t_l)\cdots f\big((2^k-1)2^{-k}\big)\sum_{l=n_1k_1}^{(n+1)k_1-1}\lambda_k(t_2)\right| \le \epsilon \sum_{l=0}^{N-1}\lambda_k(t_l).$$

Here $k_1 = 2^{N-k}$, $n_1 = 2^k$.

Since

$$(3)\qquad \sum_{k=1}^{n}\lambda_k(t_l) = tr(x(t_{l+1})-x(t_l)),$$

we have

$$(4)\qquad \sum_{l=0}^{N-1}\left(\sum_{k=1}^{n}\lambda_k(t_l)\right) = \sum_{l=0}^{N-1} tr\left(x(t_{l+1}) - x(t_l)\right) = tr\left(x(1) - x(0)\right).$$

This result, combined with the non-negativity of the $\lambda_k(t_l)$, enables us to conclude that the right-hand side of (2) is bounded by $\epsilon\ tr\left(x(1) - x(0)\right)$.

Consequently, if we show that every sum of the form $\sum_{l=M}^{M'}\lambda_k(t_l)$ converges as $N \to \infty$, it will follow that S_N converges as $N \to \infty$.

In order to establish the convergence of these sums, we shall consider the auxiliary sums

$$(5)\qquad \Sigma^{(k)} = \sum_{l=M}^{M'}\left(\sum_{S=1}^{k}\lambda_S(t_l)\right),$$

for $k = 1, 2, \cdots, n$.

§4. A Theorem of Ky Fan

The result we shall employ is

<u>Theorem</u> (Ky Fan): <u>Let the characteristic values of a symmetric matrix</u> H <u>be arranged in decreasing order</u>, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. <u>For any integer</u> q, $1 \leq q \leq n$, <u>the sum</u> $\sum_{i=1}^{q} \lambda_i$ <u>is the maximum of</u> $\sum_{j=1}^{q} (Hx_j, x_j)$ <u>where the vectors</u> x_j <u>range over all sets of</u> q <u>orthonormal vectors.</u>

The proof of this result may be found in [4].

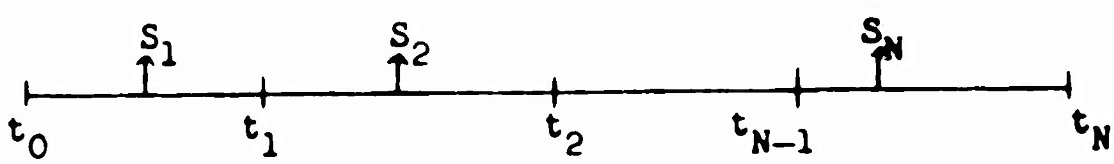
§5. <u>Continuation of the Proof</u>

Using the result stated in §4, we wish to show that

$$(1) \qquad \Sigma_2^{(k)} \leq \cdots \leq \Sigma_N^{(k)} \leq \Sigma_{N+1}^{(k)} \cdots$$

This monotonicity, taken together with the uniform boundedness of the sums, cf. (2.4), establishes convergence. Without loss of generality we may assume that the t-interval is [0, 1] in any particular sum we consider.

Let $[t_0, t_1, \cdots, t_N]$ be the set of points constituting the N^{th} subdivision, and let $S_1, S_2, \cdots S_N$ be the additional points inserted at the $(N + 1)^{st}$ subdivision, as below:



Using the representation furnished by Ky Fan's result, let us write, for a point t_i in the N^{th} subdivision,

$$(2) \qquad \sum_{S=1}^{k} \lambda_S^{(N)}(t_i) = \underset{\{x\}}{\mathrm{Max}} \sum_{j=1}^{q} \left([x(t_{i+1}) - x(t_i)]y, y \right),$$

and for the points S_{i+1} and t_i of the $(N+1)$st division

$$(3) \qquad \sum_{S=1}^{k} \lambda_S^{(N+1)}(S_{i+1}) = \operatorname*{Max}_{\{x\}} \sum_{j=1}^{q} \left([x(t_{i+1}) - x(S_{i+1})]y,\ y \right),$$

$$\sum_{S=1}^{k} \lambda_S^{(N+1)}(t_i) = \operatorname*{Max}_{\{x\}} \sum_{j=1}^{q} \left([x(S_{i+1}) - x(t_i)]y,\ y \right).$$

Since

$$(4) \qquad \left([x(t_{i+1}) - x(t_i)]y,\ y \right) = \left([x(t_{i+1}) - x(s_{i+1})]y,\ y \right) + \left([x(S_{i+1}) - x(t_i)]y,\ y \right),$$

and $\operatorname{Max}(u + v) \leq \operatorname{Max} u + \operatorname{Max} v$, we see that

$$(5) \qquad \sum_{S=1}^{k} \lambda_S^{(N)}(t_i) \leq \sum_{S=1}^{k} \lambda_S^{(N+1)}(S_{i+1}) + \sum_{S=1}^{k} \lambda_S^{(N+1)}(t_i).$$

This demonstrates the required monotonicity and completes the proof.

§6. Second Generalization

For the remainder of the paper let us assume that $x(t)$ is continuous as well as monotone increasing. We now wish to consider matrix sums of the form

$$(1) \qquad S_N = \sum_{i=0}^{N-1} \sqrt{x(t_{i+1}) - x(t_i)}\ F(t_i) \sqrt{x(t_{i+1}) - x(t_i)},$$

where $F(t)$ is a continuous matrix function over $[0, 1]$, and $\sqrt{x(t_{i+1}) - x(t_i)}$ is the unique non-negative definite square root of $x(t_{i+1}) - x(t_i)$.

The motivation for this generalized Stieltjes sum may be found in [2], where a generalization of scalar probability distributions and Markoff transformations may be found.

We conjecture the following result:

Theorem (conjecture). *As $N \to \infty$, S_N converges to a linear matrix functional which we may write* $\int_0^1 \sqrt{dx}\ \ F(t) \sqrt{dx}$.

We can prove

Theorem 2. *The above statement is true for* 2 x 2 *matrices.*

§7. *Proof of Theorem 2.*

Since $x(t) - x(S)$ is a symmetric matrix, whose elements are continuous functions of t and S, we may write it in the form

$$(1) \qquad x(t) - x(S) = T \begin{pmatrix} \lambda_1(t, S) & 0 \\ 0 & \lambda_2(t, S) \end{pmatrix} T',$$

where T is an orthogonal matrix whose elements are continuous functions of t and S, for $0 \leq S, t \leq 1$.

Furthermore, for $t \geq S$,

$$(2) \qquad \sqrt{x(t) - x(S)} = T \begin{pmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \end{pmatrix} T'.$$

Consequently, we may write

$$(3) \qquad S_N = \sum_{i=0}^{N-1} T \begin{pmatrix} \sqrt{\lambda_1(t_i)} & 0 \\ 0 & \sqrt{\lambda_2(t_i)} \end{pmatrix} (T'F(t_i)T) \begin{pmatrix} \sqrt{\lambda_1(t_i)} & 0 \\ 0 & \sqrt{\lambda_2(t_i)} \end{pmatrix} T'$$

As above, it is easy to demonstrate that the convergence of this matrix sum is equivalent to the convergence of the sum

(4) $$\Sigma_N = \sum_{i=0}^{N-1} \begin{pmatrix} \sqrt{\lambda_1(t_i)} & 0 \\ 0 & \sqrt{\lambda_2(t_i)} \end{pmatrix} C \begin{pmatrix} \sqrt{\lambda_1(t_i)} & 0 \\ 0 & \sqrt{\lambda_2(t_i)} \end{pmatrix},$$

where C is a constant matrix. The sums that arise are $\sum_i \lambda_1(t_i)$, $\sum_i \lambda_2(t_i)$, which we have already treated, and a new sum,

(5) $$G_N = \sum_{i=0}^{N-1} \sqrt{\lambda_1(t_i)\,\lambda_2(t_i)} = \sum_{i=0}^{N-1} |x(t_{i+1}) - x(t_i)|^{1/2}.$$

To establish the convergence of this sum we shall employ the same type of monotonicity argument utilized above. We require

<u>Lemma</u>. <u>Let</u> A <u>and</u> B <u>be</u> 2 x 2 <u>non-negative definite symmetric matrices. Then</u>

(6) $$\sqrt{|A + B|} \geq \sqrt{|A|} + \sqrt{|B|}.$$

<u>Proof</u>: For the 2 x 2 case, the simplest proof is computational.

Let

(7) $$A = \begin{pmatrix} a_1 & a_2 \\ a_2 & a_3 \end{pmatrix}, \quad B = \begin{pmatrix} b_1 & 0 \\ 0 & b_3 \end{pmatrix}.$$

It is easy to see that we may take one matrix in diagonal form.

Then

(8) $$\begin{aligned} |A + B| &= (a_1 + b_1)(a_3 + b_3) - a_2^2, \\ |A| &= a_1 a_3 - a_2^2 \\ |B| &= b_1 b_3. \end{aligned}$$

From this, we see that

$$(9) \qquad |A + B| = (a_1a_3 - a_2^2) + (b_1a_3 + a_1b_3) + b_1b_3 \geq (a_1a_3 - a_2^2) + b_1b_3 + 2\sqrt{b_1b_3}\sqrt{a_1a_3 - a_2^2}$$

is a consequence of

$$(10) \qquad (b_1a_3 + a_1b_3)^2 \geq 4b_1b_3(a_1a_3 - a_2^2),$$

or

$$(11) \qquad (b_1a_3 - a_1b_3)^2 + 4b_1b_3a_2^2 \geq 0.$$

Using this lemma, the proof proceeds as above.

§8. Discussion

It is seen from the foregoing that the proof of the general case rests upon establishing the convergence of sums of the form

$$(1) \qquad S_N^{(jk)} = \sum_{i=0}^{N-1} \sqrt{\lambda_j(t_i)\,\lambda_k(t_i)}\ .$$

It is not clear how one can use the previous methods to treat this general case.

Bibliography

1. Bellman, R., On a Generalization of Classical Probability Theory—I. Markoff Chains, "Proc. Nat. Acad. Sci.", Vol. 39 (1953), pp. 1075–77.

2. __________, Limit Theorems for Non-Commutative Operations—I, "Duke Math. Jour.", Vol. 21, (1954), pp. 491-500.

3. Bochner, S., Completely Monotone Functions in Partially Ordered Spaced, "Duke Math. Jour.", Vol. 9 (1942), pp. 519–526.

4. Ky Fan, On a Theorem of Weyl Concerning Eigenvalues of Linear Transformations—I, "Proc. Nat. Acad. Sci.", Vol. 35 (1949), pp. 652–55.