THE DIFFERENTIABILITY OF TRANSITION FUNCTIONS[1]

BY DONALD ORNSTEIN

Communicated by J. L. Doob, October 6, 1959

In this paper we prove that the transition functions of a denumerable Markoff chain are differentiable or equivalently: Given a matrix of real valued functions $P_{ij}(t)$ $(i, j=1, 2, \cdots)$ $0 \leqq t < \infty$ satisfying

(1) $P_{ij}(t)$ is non-negative and continuous,

$$P_{ij}(0) = \begin{cases} 1 \text{ if } i = j, \\ 0 \text{ if } i \neq j, \end{cases} \tag{2}$$

$$P_{ij}(t_1 + t_2) = \sum_{k=1}^{\infty} P_{ik}(t_1) P_{kj}(t_2), \tag{3}$$

$$\sum_{j=1}^{\infty} P_{ij}(t) = 1.^2 \tag{4}$$

Our theorem is that $P_{ij}(t)$ has a finite continuous derivative for all $t > 0$.

This result was conjectured by Kolmogoroff in [**4**].

Doob showed [**3**] that $P_{ij}(t)$ has a right hand derivative (possibly infinite) at $t=0$ and Kolmogoroff showed [**4**] that this derivative is finite if $i \neq j$, (if $i=j$ there are examples where it is infinite). Austin [**1**; **2**] showed that that $P_{ij}(t)$ has a finite continuous derivative for $t > 0$ if either $P_{ii}(t)$ or $P_{jj}(t)$ has a finite derivative at 0.

We will now give the proof[3] of our theorem. We will think of the matrices $\{P_{ij}(t)\}$ as transformations on sequences in such a way that $\{P_{ij}(t)\}$ transforms the sequence with 1 in the mth place and 0 elsewhere into the sequence whose kth term is $P_{mk}(t)$. We will use letters like v to denote a sequence, T to denote a particular matrix and $T(v)$ to denote the sequence v transformed by the matrix T.

Our first step will be to show that $P_{11}(t)$ has bounded variation in some interval (say from 0 to t_0). To do this we will estimate $\sum_{i=0}^{N-1} |P_{11}(it_0/N) - P_{11}((i+1)t_0/N)|$ for a fixed integer N. The estimate will turn out to be independent of N. To simplify notation we will let $T = \{P_{ij}(t_0/N)\}$ and let $f_i = P_{11}(it_0/N)$.

We will first define a sequence of vectors (or sequences) v_i. v_0 will be the sequence with 1 in the first place and 0 elsewhere. Let us de-

---

[1] This research was supported in part by the United States Air Force under contract AF49(638)-265, monitored by the Office of Scientific Research.

[2] This condition has been eliminated by Professor Jurkat.

[3] I wish to thank Professor Chung for valuable aid and encouragement in the preparation of this paper.

denote by v^* the sequence whose first term is 0 and which agrees with v everywhere else. Define $v_{i+1}=(T(v_i))^*$. We then have

$$T^s(v_0) = \sum_{i=0}^{s} f_{s-i}v_i. \tag{1}$$

This is easily verified by induction (note that the first coordinate of $T^s(v_0)=f_s$ by definition). We will define a sequence of positive real numbers β_i. $\beta_0=1-f_1$ and β_i $(i\geqq 1)$ is the first coordinate of $T(v_i)$. The following formula is also easy to check.

$$f_{s+1} - f_s = -f_s\beta_0 + \sum_{i=1}^{s} f_{s-i}\beta_i. \tag{2}$$

(We must interpret $\sum_{i=1}^{0}$ as 0). Rewriting (2) we get

$$f_{s+1} - f_s = f_s\sum_{i=1}^{s}\beta_i - f_s\beta_0 + \sum_{i=1}^{s}(f_{s-i} - f_s)\beta_i.$$

$$\sum_{s=0}^{N-1} |f_s - f_{s+1}| \leqq \sum_{s=0}^{N-1} \left|f_s\sum_{i=1}^{s}\beta_i - f_s\beta_0\right| + \sum_{s=0}^{N-1}\sum_{i=1}^{s}|f_{s-i} - f_s|\,\beta_i. \tag{3}$$

$$\sum_{s=0}^{N-1}\sum_{i=1}^{s}|f_{s-i} - f_s|\,\beta_i \leqq \left(\sum_{s=0}^{N-1}|f_s - f_{s+1}|\right)\left(\sum_{i=1}^{N-1} i\beta_i\right). \tag{4}$$

To see (4) note that

$$\sum_{s=0}^{N-1}\sum_{i=1}^{s}|f_{s-i} - f_s|\,\beta_i = \sum_{j=1}^{N-1}\sum_{k=j}^{N-1}|f_{k-j} - f_k|\,\beta_j$$

and

$$\sum_{k=i}^{N-1}|f_{k-j} - f_k| \leqq i\sum_{s=0}^{N-1}|f_s - f_{s+1}|.$$

From (3) and (4) we get

$$\begin{aligned}\sum_{s=0}^{N-1}|f_s - f_{s+1}| &\leqq \left(\sum_{s=0}^{N-1}|f_s - f_{s+1}|\right)\left(\sum_{i=1}^{N-1} i\beta_i\right)\\ &\quad + \sum_{s=0}^{N-1}\left|f_s\sum_{i=1}^{s}\beta_i - f_s\beta_0\right|.\end{aligned} \tag{5}$$

If we now assume that the t_0 we used in defining T has the property that $P_{11}(t)>3/4$ for all $t<t_0$ we will be able to show that both $\sum_{i=1}^{N-1} i\beta_i$ and $\sum_{s=0}^{N-1}\left|f_s\sum_{i=1}^{s}\beta_i-f_s\beta_0\right|$ are $<1/2$. This and (5) will then immediately imply that $\sum_{s=0}^{N-1}|f_s-f_{s+1}|<1$ and, since $P_{11}(t)$ is continuous and our estimate does not depend on N, that the variation of $P_{11}(t)$ $(t<t_0)$ is $\leqq 1$. To get $\sum_{i=1}^{N-1} i\beta_i < 1/2$ we note first that $\sum_{i=1}^{N-1} i\beta_i < \sum_{i=1}^{N} |v_i|$ ($|v|$ = sum of the absolute values of the coordinates of v) since $\beta_i = |v_i| - |v_{i+1}|$. Next we show that $\sum_{i=1}^{N} |v_i| < 1/2$. $T^N(v_0) = f_N v_0 + \sum_{i=1}^{N} f_{N-i} v_i$ and since row sums $= 1$, $\sum_{i=1}^{N} f_{N-i}|v_i| = 1 - f_N < 1/4$. Each of the $f_{N-i} > 1/2$ so $\sum_{i=1}^{N}|v_i| < 1/2$. $\sum_{s=0}^{N-1} |f_s \sum_{i=1}^{s} \beta_i - f_s\beta_0| < 1/2$ because $|\beta_0 - \sum_{i=1}^{s} \beta_i| = |v_{s+1}|$.

We now know that $P_{11}(t)$ has variation <1 in a certain interval about 0. The following argument shows that the variation of $P_{1j}(t) \leqq 4$ in the same interval.

$$T^{s+1}(v_0) - T^s(v_0) = \sum_{i=0}^{s+1} (f_{s+1-i} - f_{s-i})v_i, \qquad (f_{-1} = 0)$$

$$\text{(6)} \qquad \sum_{s=0}^{N-1} |T^{s+1}(v_0) - T^s(v_0)| \leqq \sum_{i=0}^{N} \sum_{s=i-1}^{N-1} |(f_{s+1-i} - f_{s-i})v_i| \leqq \sum_{i=0}^{N} 2|v_i| \leqq 4.$$

The remainder of the proof follows a suggestion of K. L. Chung.[4] Functions of bounded variation have a finite derivative almost everywhere and we can therefore pick a $t_1 < t_0$ such that $P_{1j}(t)$ has a derivative at t_1 for all j. For an arbitrary t_2 the existence of a derivative for $P_{1i}(t_1+t_2)$ $(i = 1 \cdots \infty)$ follows from the fact that

$$\frac{P_{1i}(t_1 + t_2) - P_{1i}(t_1 + t_2 + \alpha)}{\alpha} = \sum_{k=1}^{\infty} \frac{P_{1k}(t_1) - P_{1k}(t_1 + \alpha)}{\alpha} P_{ki}(t_2)$$

and the following lemma: given ϵ there exists an integer K such that

$$\text{(7)} \qquad \sum_{j=K}^{\infty} \frac{|P_{1j}(t_1) - P_{1j}(t_1 + \alpha)|}{\alpha} < \epsilon, \qquad \frac{t_1}{4} > \alpha > 0.$$

We conclude by proving (7). For a given $\alpha < t_1/4$ we will pick a t_0' between t_1 and $t_1/2$ and an integer N such that $t_0'/N = \alpha$ and we will define T and v_i as before, except that we will use t_0' instead of t_0.

It is easy to show that given ϵ_1 (we will pick ϵ_1 to be $<(1/8)\epsilon \cdot t_1/2 \cdot 1/2$) there is a K_1 such that $\sum_{j=K_1}^{\infty} P_{1j}(t) < \epsilon_1$ for all $t < t_1$. We then have $\sum_{i=1}^{N} |v_i^{K_1}| < 2\epsilon_1$ ($|v_i^{K_1}|$ is the sum of the absolute values of the terms of v_i with index $\geqq K_1$). The same argument as the one used in (6) shows

[4] The original proof did not make use of the theorem that functions of bounded variation have derivatives almost everywhere and was very much longer. Professor Chung's idea also gives $P'_{1i}(t_1+t_2) = \sum_k P_{1k}(t_1)P'_{ki}(t_2)$. Professor Chung has also proved (in a different way) that $P'_{1i}(t_1+t_2) = \sum_k P_{1k}(t_1)P'_{ki}(t_2)$.

$$\sum_{s=1}^{N-1} \sum_{j=K_1}^{\infty} | P_{1j}((s+1)\alpha) - P_{1j}(s\alpha) | < 4\epsilon_1. \tag{8}$$

There are at least $(N-1)/2$ integers s such that

$$\sum_{j=K_1}^{\infty} | P_{1j}((s+1)\alpha) - P_{1j}(s\alpha) | < 8\epsilon_1 \frac{1}{N} \tag{9}$$

and for one of these, call it r,

$$\sum_{j=1}^{K_1} | P_{1j}((r+1)\alpha) - P_{1j}(r\alpha) | < \frac{8}{N} \cdot \tag{10}$$

This follows from (6). We now pick ϵ_2 (make it $<\epsilon\cdot(1/8)K_1\cdot t_{1/2}\cdot 1/2$). There is a $K>K_1$ such that

$$\sum_{j=K}^{\infty} P_{ij}(t) < \epsilon_2 \quad \text{for all } t < t_1 \text{ and } i \leqq K_1,$$

$$\begin{aligned}
&\sum_{j=K}^{\infty} | P_{1j}(t_1) - P_{1j}(t_1+\alpha) | \\
&\qquad \leqq \sum_{m=K}^{\infty} \sum_{j=1}^{\infty} | P_{1j}(r\alpha) - P_{1j}(r+1)\alpha | P_{jm}(t_1 - r\alpha) \\
&\qquad = \sum_{m=K}^{\infty} \sum_{j=K_1+1}^{\infty} | P_{1j}(r\alpha) - P_{1j}((r+1)\alpha) | P_{jm}(t - r\alpha) \\
&\qquad\quad + \sum_{m=K}^{\infty} \sum_{j=1}^{K_1} | P_{1j}(r\alpha) - P_{1j}((r+1)\alpha) | P_{jm}(t_1 - r\alpha).
\end{aligned} \tag{11}$$

The first term of this last expression is $< 8\epsilon_1\cdot 1/N$ by (9), $|P_{1j}(r\alpha)-P_{1j}((r+1)\alpha| <8/N$ by (10) and $\sum_{m=K}^{\infty} P_{jm}(t_1-r\alpha)<\epsilon_2$ for each $j<K_1$ by (11). Hence the second term is $<8/N\cdot\epsilon_2\cdot K_1$.

This finishes the proof of the lemma.

Bibliography

1. D. G. Austin, *On the existence of the derivative of Markoff transition probability functions*, Proc. Nat. Acad. Sci. U. S. A. vol. 41 (1941) pp. 224–226.

2. ———, *Some differentiability properties of Markoff transition probability functions*, Proc. Amer. Math. Soc. vol. 7 (1956) pp. 751–761.

3. J. L. Doob, *Topics in the theory of Markoff chains*, Trans. Amer. Math. Soc. vol. 52 (1942) pp. 37–64.

4. A. N. Kolmogoroff, *On the differentiability of transition probabilities in stationary Markov processes with a denumerable number of states*, Moskov. Gos. Univ. Uč. Zap. Mat. vol. 148 (1951) pp. 53–59 (Russian).

University of Wisconsin