Commentationes Mathematicae Universitatis Carolinae

Eugene Seneta

The principle of truncations in applied probability

Commentationes Mathematicae Universitatis Carolinae, Vol. 9 (1968), No. 2, 237--242

Persistent URL: http://dml.cz/dmlcz/105175

Commentationes Mathematicae Universitatis Carolinae
9, 2 (1968)

THE PRINCIPLES OF TRUNCATIONS IN APPLIED PROBABILITY +

E. SENETA, Canberra

1. The Principle of Truncations

The "principle of truncations" in general may be stated as follows: "To solve an infinite system of linear equations in an infinite number of unknowns, one limits the system to the first n equations, and in it neglects all but the first n unknowns. The solutions of these systems tend, as to the required solution". A history of this principle up to 1913 is contained in Chs. I and II of Riesz (1913), from which work the contents of the present section are taken.

The use of the idea appears to go right back to Fourier (The Analytical Theory of Heat) who obtained a correct answer, as did several subsequent workers, without actual justification. The legitimacy of the method was first investigated by Poincaré, who showed that if the matrix A and the infinite vector c of the system under consideration

$$Ax = c \ (\text{i.e.} \sum_{k=1}^{\infty} a_{ik} x_k = c_i \,, \ i = 1, 2, \dots) \tag{1.1}$$

satisfy extremely strict conditions, the principle is valid.

The problem is closely connected to the problem of in-

+ An expository article, read at the 1967 meeting of the Australian Mathematical Society, in Canberra.

finite determinants. The solution $y(n)$ of the $n \times n$ truncated system

(1.2) $\sum_{k=1}^{n} a_{ik} y_k(n) = c_i$ is given by $y_k(n) = \Delta_n^{(k)}/\Delta_n$

$i, k = 1, \dots, n$, if $\Delta_n \neq 0$, using the usual determinantal notation. Supposing $\Delta_n \neq 0$ for all n sufficiently large, according to the principle:

$x_k = \{\lim n \to \infty\, y_k(n)\}$, supposing the latter limit exists. However it is easy to find simple systems where <u>both</u> conditions <u>hold</u>, but the last equality <u>does not</u>.

For matrices A of a certain type viz.

(1.3) $a_{ij} = b_{ij} + \delta_{ij}$, $\sum_{i,j} |b_{ij}| < \infty$, ($\delta_{ij}$ = Kronecker δ)

most of the desirable properties hold e.g. $\Delta = \lim n \to \infty\, \Delta_n$ exists.

Such determinants are called <u>normal</u>, and have many of the properties of finite determinants e.g. row-column expansion by "cofactors". Moreover, if in this "normal" case (1.3) one restricts oneself in (1.1) to both c and x uniformly bounded elementwise, the usual properties of (finite) linear systems hold. Thus if $\Delta \neq 0$, the solution x is unique and given by $x_k = \Delta^{(k)} / \Delta$. ($\Delta^{(k)} = \lim n \to \infty\, \Delta_n^{(k)}$ exists.)

The most prolific worker in infinite determinants was Von Koch, who generalized normal determinants somewhat to <u>absolutely convergent determinants</u>. His work was virtually all done by the time the book of Riesz appeared. This work heralded the end of research on infinite determinants by its assertion at the beginning of Ch.3, that the heavy restrictions imposed on the matrix A in order to utilize infinite determinants were necessitated not by the problem but by the method. The aim of Riesz was to consider the existence of solutions of (1.1) rather than their explicit value. As is well known this idea led to his book becoming one of the early milestones of functional analysis.

2. Non-Negative Matrices

As the theory of infinite determinants was coming to a halt, two new branches of mathematics were evolving, which were to become intimately connected.

Perron and Frobenius, in Germany, were developing the spectral theory of (finite square) matrices with non-negative entries. In Russia, A.A. Markov was initiating the study of Markov chain, which is tantamount to the study of square stochastic matrices $P = \{p_{ij}\}$, $p_{ij} \geq 0$, $\sum_j p_{ij} = 1$.

Clearly, if the Markov chain is finite, i.e. P is finite, then the Perron-Frobenius theory is applicable. This approach to the treatment of finite Markov chains was exploited by Fréchet (1938) and then V.I. Romanovsky in particular, who wrote a large number of papers, and a treatise (1949) of over 430 pages on it. The Perron-Frobenius theory is presented and applied to Markov chains also in the book of Gantmacher (1959).

However, it is clear that finite matrix spectral methods are insufficient to cover the general theory of Markov chains, where P may be infinite. Thus these methods, with one notable exception, were set aside in the post war years, and the general theory of Markov chains was developed by a generating function approach, initiated by Feller in his famous book of 1951. Chung's monograph of 1960 reviewed the theory from this viewpoint. There were also matrix approaches, but of a rather different kind, to finite chains e.g. Kemeny and Snell (1960), Howard (1960).

The one exception (to the general tendency) was a part (22-24) of the work of Sarymsakov (1954), who tried to apply the principle of truncations (still using cofactor-determinant methods) to infinite stochastic matrices, with the purpose of providing a calculational algorithm for various quantities of interest. However, he ran into trouble, as had earlier workers in infinite determinants, because of his method.

A more recent development in a new direction was the work of Vere-Jones (1962), who applied generating function methods to (infinite)non-negative matrices,satisfying mild restrictions, to obtain results very analogous to the Perron-Frobenius theory.

By considering the approach of Vere-Jones in conjunction with the general approach of Sarymsakov, the tool for investigating the validity of the method of truncations for infinite Markov chains had become available. Moreover, it had become important to do this, for generating functions are of no use in a practical sense, and hence some approximation method (of truncation) is desirable. In fact the wheel has come full circle, and we are often interested in the actual solutions of equations such as (1.1), under conditions which, a priori, determine their existence and uniqueness.

3. Infinite Denumerable Markov Chains

The Markov chains of most interest are irreducible i.e. for every two indices i, j there exists a finite sequence $k_1, k_2, \ldots k_n$, such that $p_{ik_1} p_{k_1 k_2} \cdots p_{k_n j} > 0$. These, moreover, fall into two classes: recurrent and transient chains. Recurrent chains have a unique positive (upto multiplicative const.) solution x (the invariant measure) of the equations

$$x' P = x' \quad \text{i.e.} \quad x'(I - P) = 0' .$$

We note that (without some adjustment, possibly) the matrix $A = I - P$ never satisfies the condition (1.3) for a normal determinant.

However, the method of truncations "works", insofar as the following result is valid. (By $(n)C_{ji}$ denote the cofactor of the (j, i) element of the $n \times n$ truncation $[_{(n)}I - _{(n)}P]$ of $I - P$, and put $\Delta_n = \det[_{(n)}I - _{(n)}P]$).

Theorem: (A) As $n \to \infty$, $(n)C_{ji} / (n)C_{ii} \uparrow V_{ji}$ for all i, j. For fixed i, $\{V_{ji}\}$ is the unique positive invariant measure of P with $V_{ii} = 1$ if the Markov chain is recurrent.

(B) $\Delta_n \downarrow \Delta \geq 0$ as $n \to \infty$. $\Delta > 0$ implies the chain is transient.

(C) $\lim_{n \to \infty} (n)C_{ji} = C_{ji}$ exists and is (finite and) non-negative for all i, j. If $C_{ji} > 0$ for some i, j, this holds true for all i, j.

The proof depends largely on the techniques of Vere-Jones and Sarymsakov, and may be found in Seneta (1967) together with results of like nature. (Actually, a more complete proof will appear in a sequel to this paper, in the same journal.)

References

[1] M. FRÉCHET (1938(: Théorie des événements en chaine dans le cas d'un nombre fini d'états possible. Gauthier Villars.

[2] F.R. GANTMACHER: (1959): Applications of the Theory of Matrices. Interscience.

[3] R.A. HOWARD(1960(: Dynamic Programming and Markov Processes. Wiley.

[4] J.G. KEMENY - SNELL L.J.(1960): Finite Markov Chains. Van Nostrand.

[5] F. RIESZ (1913): Les Systèmes d'Equations Linéaires a une Infinite-d'Inconnues. Gauthier-Villars.

[6] V.I. ROMANOVSKY (1949): Discrete Markov Chains. (Russian) GITTL.

[7] T.A. SARYMSAKOV (1954): Elements of the theory of Markov Processes. (Russian) GITTL.

[8] E. SENETA (1967): Finite approximations to infinite non-negative matrices. Proc. Camb. Phil. Soc. 63, 983-992. x)

[9] D. VERE-JONES (1962): Geometric ergodicity in denumerable Markov chains. Quart. J. Math. Oxford Ser. 13, 7-28.

(Received March 20,1968)

---

x) A more complete reference list is given in this article.