# —NOTES—

## ON THE CONVERGENCE OF MATRIX ITERATION PROCESSES*

By ROBERT PLUNKETT (*The Rice Institute*)

A recent paper [1]† has presented a formal proof of an iteration method for the solution of non-homogeneous (Fredholm) integral equations which is directly applicable to non-homogeneous matrix equations. Various special cases of this method have been used for matrix equations (cf. [2], [3]) but no general statement has yet appeared. The proof in the case of matrices is much simpler and the conditions for convergence are less stringent than in the case of integral equations.

The equation under consideration is

$$u = u_0 + Qu\,, \tag{1}$$

where $u$ is an unknown column vector of $n$ elements, $u_0$ is a known column vector and $Q$ is a square matrix of $n^2$ elements. The formal solution is obviously

$$u = (I - Q)^{-1}u_0\,, \tag{1a}$$

but the reciprocal matrix is not always easily determined if $n$ is very large. Associated with Eq. (1) is the homogeneous equation

$$\lambda_r u_r = Qu_r\,, \tag{2}$$

in which $\lambda_r$ are the latent roots and $u_r$ , the corresponding latent vectors.

The usual iteration method is to let

$$v_h = u_0 + Qv_{h-1} \tag{3}$$

where $v_0 = u + w$ is the first approximation to the answer and $v_h$ is the $h$th iterated approximation. For rapid convergence, an attempt is made to make $w$ as small as possible. Then

$$\begin{aligned} v_1 &= u_0 + Qv_0 = u + Qw, \\ v_h &= u + Q^h w. \end{aligned} \tag{3a}$$

For a matrix $Q$ with $n$ distinct, non-repeated roots, Eq. (3a) is easily expanded into a simpler form. When all the roots are distinct, all the latent vectors are linearly independent, and $w$ may be expressed uniquely as a linear combination of them:

$$w = \sum_{r=1}^{n} A_r u_r\,,$$

where $A_r$ are scalars, and it is obvious that

$$Q^h w = \sum_{r=1}^{n} \lambda_r^h A_r u_r\,, \tag{3b}$$

which must converge to zero for all $|\lambda_r| < 1$.

*Received Feb. 21, 1949.

†Numbers in brackets refer to the bibliography at the end of this paper.

If the roots are not all distinct, let us assume that there is one $\lambda_s$ which is repeated $s$ times, and that $| I\lambda_s - Q |$ is $q$ times degenerate, i.e., that there are $q$ linearly independent vectors associated with this root. If $q = s$, the equation may be treated as in Eq. (3b). If $q \neq s$, the confluent form of Sylvester's Theorem is applicable and states (cf. [4]) that

$$P(Q) = \sum T(\lambda_s), \tag{3c}$$

the summation to be taken over all distinct values of $\lambda_s$, where $P(Q)$ is any polynomial in $Q$ and

$$T(\lambda_s) = P(\lambda_s)Z_{s-1}(\lambda_s) + \frac{P^{(1)}(\lambda_s)Z_{s-2}(\lambda_s)}{1!} + \cdots + \frac{P^{(s-1)}(\lambda_s)Z_0(\lambda_s)}{(s-1)!},$$

$$Z_i(\lambda_s) = \frac{1}{i!}\left[\frac{d^i}{d\lambda^i}\left(\frac{F(\lambda)}{\Delta_s(\lambda)}\right)\right]_{\lambda=\lambda_s}.$$

Now $P^{(i)}(\lambda_s)$ is the $i$th derivative of $P$ with respect to $\lambda$ at $\lambda_s$, $F(\lambda)$ is the adjoint of $(I\lambda - Q)$, and

$$\Delta_s(\lambda) = (\lambda - \lambda_{s+1})(\lambda - \lambda_{s+2}) \cdots (\lambda - \lambda_n),$$

i.e., the products of the differences involving all the roots except the $s$ repeated roots under consideration, but including the repeated values of all other repeated roots. Letting

$$P(Q) = Q^h = \sum T_h(\lambda_s),$$

then

$$T_h(\lambda_s) = \lambda_s^h Z_{s-1}(\lambda_s) + \frac{h\lambda_s^{h-1}}{1!} Z_{s-2}(\lambda_s) + \cdots + \frac{h(h-1)\cdots(h-s+2)\lambda^{h-s+1}}{(s-1)!} Z_0(\lambda_s),$$

$$T_{h+1}(\lambda_s) = \lambda_s^{h+1}Z_{s-1}(\lambda_s) + \frac{(h+1)\lambda_s^h}{1!} Z_{s-2}(\lambda_s)$$

$$+ \cdots + \frac{(h+1)h\cdots(h-s+3)\lambda^{h-s+2}}{(s-1)!} Z_0(\lambda_s);$$

and for large $h$ it is apparent that

$$Q^{h+1}w = \sum T_{h+1}(\lambda_s)w = \sum \lambda_s T_h(\lambda_s)w,$$

since the matrices $Z$ do not depend on $h$. The validity of this last expression is increased for multiple degeneracy of $(I\lambda_s - Q)$; if the degeneracy is $q$, $F(\lambda_s)$ and all of its derivatives up to the $(q-2)$th are zero, which means that

$$Z_i(\lambda_s) = 0 \qquad (i \leq q - 2)$$

or that the last $q - 1$ terms in $T(\lambda_s)$ vanish. This is to be expected, since for degeneracy $s$, Eq. (3b) holds exactly. Then for large $h$, Eq. (3a) may be written

$$Q^h w = \sum \lambda_r^h B_r w, \tag{3d}$$

where $B_r$ is a matrix involving the $Z_i(\lambda_s)$ and is independent of $h$ in the limit, and the summation is taken over all distinct values of $\lambda_r$. Thus it may be seen that Eq. (3a) converges to $u$ for large $h$ in all cases, if all $| \lambda_s | < 1$.

The proposed solution is to let

$$v_h = \theta v_{h-1} + (1 - \theta)[u_0 + Qv_{h-1}], \tag{4}$$

where

$$v_1 = \theta v_0 + [u_0 + Qv_0](1 - \theta)$$

$$= u + [\theta I + (1 - \theta)Q]w$$

or

$$v_h = u + [\theta I + (1 - \theta)Q]^h w. \tag{4a}$$

Comparing this with Eq. (3a), it may be seen that the convergence depends on the roots of $[\theta I + (1 - \theta)Q]$ which are $[\theta + (1 - \theta)\lambda_r]$. Thus if all

$$| \theta + (1 - \theta)\lambda_r | < 1, \tag{4b}$$

Eq. (4a) will converge to $u$.

There are cases in which it is impossible to find a number $\theta$ which will satisfy Eq. (4b) for all $\lambda_r$, without manipulating the Eqs. (1) so as to change $Q$ and thus the roots. For example, if there are two of the roots which are real, such that $0 < \lambda_p < 1 < \lambda_q$, then for $\lambda_p$, $\theta < 1$ and for $\lambda_q$, $\theta > 1$ which is impossible. But if all $\lambda_r$ are real and less than 1, where $\lambda_p$ is the largest negative root, $(|\lambda_p| - 1)/(|\lambda_p| + 1) < \theta < 1$. Similarly, for all $\lambda_r$ real and greater than 1, $1 < \theta < (\lambda_p + 1)/(\lambda_p - 1)$, where $\lambda_p$ is the largest positive root. While the exact value of $\theta$ within the limits is not important, it may be seen that all values of $| \theta + (1 - \theta)\lambda_r |$ should be as small as possible for best convergence. For example, in one actual case [2] there was one root of $-2$, two near $-1$ and several negative and very small, $\theta$ was taken to be 1/2, the largest value of $| \theta + (1 - \theta)\lambda_r |$ was about 1/2 and five iterations reduced the error to $(1/2)^5$ or about 3%.

If all the $\lambda_r$ are real, positive and less than one, this method is not necessary for convergence, but hastens it if any $\lambda_r$ are close to one.

Equation (4) is equivalent to a Toeplitz summation on the vectors $v_h$ as specified by Eq. (3), with the summation matrix (cf. [5])

$$a_{ij} = \begin{Bmatrix} i \\ j \end{Bmatrix}(1 - \theta)^{i-j}\theta^j, \qquad j \leq i,$$

$$= 0, \qquad j > i.$$

## Bibliography

[1] Hans Buckner, *A special method of successive approximations for Fredholm integral equations*, Duke Math. J. 15, 197-206 (1948).

[2] R. Plunkett, *A matrix method of calculating propeller blade moments and deflections*, J. Appl. Mech., forthcoming.

[3] R. Plunkett, *Forced vibrations of a rotating propeller*, Doctoral Thesis, M. I. T., 1948.

[4] Frazer, Duncan and Collar, *Elementary matrices*, Cambridge University Press, 1938, p. 83.

[5] A. Zygmund, *Trigonometrical series*, Monografje Matematyczne, Warsaw, 1935, p. 40.