On Nearly Triangular Matrices [1]

A. M. Ostrowski [2]

A discussion is presented of the change in the inverse of a triangular matrix if on one side the zeros are replaced by sufficiently small numbers and on the other side the non-vanishing elements are varied by sufficiently small amounts.

1. Introduction

Consider a system of linear equations

$$\sum_{\nu=1}^{n} a_{\mu\nu} x_\nu = y_\mu \qquad (\mu = 1, \ldots, n), \tag{1}$$

with the matrix A, in which all diagonal elements $a_{\mu\mu}$ ($\mu = 1, \ldots, n$) are not equal to 0 and the elements off the diagonal satisfy for two positive numbers m, M the inequalities

$$\left.\begin{array}{l} |a_{\mu\nu}| \leqq m|a_{\mu\mu}| \quad (\nu < \mu;\ \mu = 1, \ldots, n), \\[4pt] |a_{\mu\nu}| \leqq M|a_{\mu\mu}| \quad (\nu > \mu;\ \mu = 1, \ldots, n-1). \end{array}\right\} \tag{2}$$

If m is very small, the system does not essentially differ from the corresponding "triangular" system in which all $a_{\mu\nu}$ with $\nu < \mu$ are replaced by zeros and the matrix of which will be denoted by $A^{(0)}$. It then appears plausible that the solution of this triangular system does not differ very much from that of the system (1).

However, the value of the determinant of the order n

$$\begin{vmatrix} 1 & -M & 0 & & 0 & 0 \\ 0 & 1 & -M & & & 0 \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & \cdot & \\ & & & & \cdot & \\ 0 & & & & 1 & -M \\ -m & 0 & & & 0 & 1 \end{vmatrix} = 1 - mM^{n-1},$$

shows that if M is, for instance, greater than or equal to 10, the determinant of our system will not be necessarily different from zero unless $m < 10^{-(n-1)}$. A detailed study of the problems connected with the matrices characterized by (2) appears, therefore, to be of importance and interest.

As the first problem in this connection, we give a necessary and sufficient condition that any matrix A satisfying conditions (2) be nonsingular. If $m < M$, this condition is given by

$$\frac{m}{(1+m)^n} < \frac{M}{(1+M)^n}, \tag{3}$$

and, if $m = M$, by [3]

$$m < \frac{1}{n-1}.$$ (3°)

In order to obtain a precise measure of the influence of the change from A to $A^{(0)}$, we have to discuss the estimates (for convenient norms) of the norm of the matrix $A^{-1} - A^{(0)-1}$.

We consider, in particular, two such norms defined in section 5 and denoted by $|A^{-1} - A^{(0)-1}|_p$, $(p = 1, \infty)$, which are particularly suitable for the problems of numerical analysis. Assuming, without loss of generality, that $a_{\mu\mu} = 1$ $(\mu = 1, \ldots, n)$, we show that for given values of m, M such that $M \geq 1.5/n$, $n \geq 4$, we have

$$|A^{-1} - A^{(0)-1}|_p \leq (1+M)^{n-1} \frac{\delta}{1-\delta} \quad \left(M \geq \frac{1.5}{n}, \; n \geq 4 \right),$$ (4)

where $1 - \delta$ is the smallest modulus of the determinant attainable for the matrices A and is connected with m by the relation

$$\delta = \frac{M_n m}{1 - \theta \dfrac{m}{M}}, \qquad 0 < \theta < 1,$$ (5)

with

$$M_n = \frac{(1+M)^n - nM - 1}{M}.$$ (6)

If $M < 1.5/n$, the formula (4) need not be valid any longer, but we can prove in this case the relation

$$|A^{-1} - A^{(0)-1}|_p \leq \frac{6nm}{1-2nm} \qquad (M \leq 1.5/n)$$ (7)

is valid as long as m remains less than $1/2n$.

The estimate (7) is not a "best" estimate for all values of $M \leq 1.5/n$, but still it is not far from the best, since for $m = M \leq 1/(n-1)$ we have

$$|A^{-1} - A^{(0)-1}|_p \leq \frac{(n-1)m}{1 - (n-1)m} \qquad \left(m = M < \frac{1}{n-1} \right),$$ (8)

which cannot be improved for any value of $m < 1/(n-1)$.

The condition (3) is derived in section 4, theorem B. However, we derive it as a special case of a more general theorem, where in the inequalities (2) the expressions m, M *depend on* μ, that is to say, change from one row to another. The necessary and sufficient condition for all matrices A to be regular (theorem A, section 4) is in this case rather unwieldy, but still may be very useful in some cases because it contains $2n - 2$ instead of two essential parameters. The direct derivation of theorem B is, of course, much simpler, since, as the reader will immediately see, the computations of the determinant Ω_n in section 2 can be considerably shortened in this case. The connection between the formal algebra of sections 2 and 3 and theorems A and B is provided by a result concerning the so-called H-determinants and M-determinants published 16 years ago [1]. The results about the norms $|A^{-1} - A^{(0)-1}|_p$ are obtained by using the explicit representation of the inverse matrix of a certain matrix Δ_n, which provides a majorant for all matrices A^{-1}. The formulas giving Δ_n^{-1} are derived in the second part of section 3, and in section 6 the norms $|\Delta_n^{-1} - \Delta_n^{(0)-1}|_p$ are derived and discussed. The corresponding inequalities for $|A^{-1} - A^{(0)-1}|_p$ are then obtained in section 8, by using a new theorem (lemma III) concerning the connection between the H-determinants and the M-determinants.

---

[3] This condition (3°) is already contained in some results in a paper [2][4], and also can be deduced from a well-known theorem of the theory of determinants discussed in [4].

[4] Figures in brackets indicate the literature references at the end of this paper.

In this section is given another application of this theorem in estimating the variation in the inverse matrix of a triangular matrix satisfying the conditions (2) with $m=0$. We obtain an unexpectedly simple and elegant formula (120).

In section 8, the results are applied explicitly to the problems concerning the linear system (1). It may be finally remarked that these results, with obvious changes, remain valid if in the matrix of (1) the rows and the columns are interchanged, although no mention is made of it explicitly at every step.[5]

2. Value of the Determinant Ω_n

Let K_n be defined by

$$
K_n =
\begin{vmatrix}
\kappa_n & \delta_n & 0 & 0 & \ldots & 0 & 0 \\
\kappa_{n-1} & \kappa_{n-1} & \delta_{n-1} & 0 & \ldots & 0 & 0 \\
\kappa_{n-2} & \kappa_{n-2} & \kappa_{n-2} & \delta_{n-2} & \ldots & 0 & 0 \\
\cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\
\kappa_2 & \kappa_2 & \kappa_2 & \kappa_2 & \ldots & \kappa_2 & \delta_2 \\
\kappa_1 & \kappa_1 & \kappa_1 & \kappa_1 & \ldots & \kappa_1 & \kappa_1
\end{vmatrix}
\qquad (n \geq 3),
\tag{9}
$$

$$
K_1 = \kappa_1, \qquad K_2 =
\begin{vmatrix}
\kappa_2 & \delta_2 \\
\kappa_1 & \kappa_1
\end{vmatrix}
= \kappa_1(\kappa_2 - \delta_2),
$$

where in the μ-th row all elements to the left of the main diagonal and on this diagonal are equal to $\kappa_{n-\mu+1}$, the next element to the right is $\delta_{n-\mu+1}$, and all other elements are 0. κ_μ, δ_μ are here independent variables. In subtracting the second column from the first, we obtain $K_n = (\kappa_n - \delta_n) K_{n-1}$, and therefore the following formula, valid also for $n = 1, 2$:

$$
K_n = \kappa_1 \prod_{\nu=2}^{n} (\kappa_\nu - \delta_\nu).
\tag{10}
$$

Consider now for $n \geq 3$ the determinant

$$
T_n =
\begin{vmatrix}
\delta_n & 0 & 0 & 0 & \ldots & 0 & \beta_n \\
\gamma_{n-1} & \delta_{n-1} & 0 & 0 & \ldots & 0 & \beta_{n-1} \\
\gamma_{n-2} & \gamma_{n-2} & \delta_{n-2} & 0 & \ldots & 0 & \beta_{n-2} \\
\cdot & \cdot & \cdot & \cdot & \cdot \cdot \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot \cdot \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot \cdot \cdot & \cdot \\
\gamma_2 & \gamma_2 & \gamma_2 & \gamma_2 & \ldots & \delta_2 & \beta_2 \\
1 & 1 & 1 & 1 & \ldots & 1 & 0
\end{vmatrix}
\qquad (n \geq 3).
\tag{11}
$$

[5] Some of the results contained in the sections 2 to 5 have been published without proof in [3].

We have in particular

$$T_3 = \begin{vmatrix} \delta_3 & 0 & \beta_3 \\ \gamma_2 & \delta_2 & \beta_2 \\ 1 & 1 & 0 \end{vmatrix} = -(\delta_3\beta_2 + \beta_3(\delta_2 - \gamma_2)). \tag{12}$$

Developing T_n in the elements of the first line and using the value (10) of K_n, we obtain for $n \geq 4$

$$T_n = \delta_n T_{n-1} - \beta_n \prod_{\nu=2}^{n-1}(\delta_\nu - \gamma_\nu),$$

$$\frac{T_n}{\delta_2 \ldots \delta_n} = \frac{T_{n-1}}{\delta_2 \ldots \delta_{n-1}} - \frac{\beta_n}{\delta_n} \prod_{\nu=2}^{n-1}\left(1 - \frac{\gamma_\nu}{\delta_\nu}\right),$$

and therefore generally for $n \geq 4$

$$\frac{T_n}{\delta_2 \ldots \delta_n} = -\left[\frac{\beta_n}{\delta_n} \prod_{\nu=2}^{n-1}\left(1 - \frac{\gamma_\nu}{\delta_\nu}\right) + \frac{\beta_{n-1}}{\delta_{n-1}} \prod_{\nu=2}^{n-2}\left(1 - \frac{\gamma_\nu}{\delta_\nu}\right) + \cdots + \frac{\beta_4}{\delta_4} \prod_{\nu=2}^{3}\left(1 - \frac{\gamma_\nu}{\delta_\nu}\right)\right] + \frac{T_3}{\delta_2\delta_3}. \tag{13}$$

Since by (12)

$$\frac{T_3}{\delta_2\delta_3} = -\frac{\beta_3}{\delta_3}\left(1 - \frac{\gamma_2}{\delta_2}\right) - \frac{\beta_2}{\delta_2}$$

we obtain

$$\frac{-T_n}{\delta_2 \ldots \delta_n} = \sum_{\mu=2}^{n} \frac{\beta_\mu}{\delta_\mu} \prod_{\nu=2}^{\mu-1}\left(1 - \frac{\gamma_\nu}{\delta_\nu}\right),$$

where $\prod\limits_{\nu=2}^{1}$ is identically 1, and therefore finally

$$T_n = -\sum_{\mu=2}^{n} \beta_\mu \prod_{\nu=\mu+1}^{n} \delta_\nu \prod_{\nu=2}^{\mu-1}(\delta_\nu - \gamma_\nu). \tag{14}$$

If we now put

$$T_n^* = \begin{vmatrix} \delta_1 & 0 & 0 & . & . & . & 0 & \beta_1 \\ \gamma_2 & \delta_2 & 0 & . & . & . & 0 & \beta_2 \\ . & . & . & & & & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & & & & . & . \\ \gamma_{n-1} & \gamma_{n-1} & \gamma_{n-1} & . & . & . & \gamma_{n-1} & \beta_{n-1} \\ 1 & 1 & 1 & . & . & . & 1 & 0 \end{vmatrix} \qquad (n \geq 3), \tag{15}$$

$$T_3^* = \begin{vmatrix} \delta_1 & 0 & \beta_1 \\ \gamma_2 & \delta_2 & \beta_2 \\ 1 & 1 & 0 \end{vmatrix},$$

this becomes T_n if the indices of β_ν, γ_ν, δ_ν are replaced by their complements with respect to $n+1$. We obtain then from (14)

$$T_n^* = -\sum_{\mu=2}^{n} \beta_{n+1-\mu} \prod_{\nu=\mu+1}^{n} \delta_{n+1-\nu} \prod_{\nu=2}^{\mu-1}(\delta_{n+1-\nu} - \gamma_{n+1-\nu}),$$

322

or in replacing the summation index μ by $n+1-\kappa$

$$T_n^* = -\sum_{\kappa=1}^{n-1} \beta_\kappa \prod_{\nu=n+2-\kappa}^{n} \delta_{n+1-\nu} \prod_{\nu=2}^{n-\kappa} (\delta_{n+1-\nu} - \gamma_{n+1-\nu})$$

and finally, if in both products ν is replaced by $n+1-\lambda$,

$$T_n^* = -\sum_{\kappa=1}^{n-1} \beta_\kappa \prod_{\lambda=1}^{\kappa-1} \delta_\lambda \prod_{\lambda=\kappa+1}^{n-1} (\delta_\lambda - \gamma_\lambda). \tag{16}$$

Consider now for $n \geqq 3$ the determinant

$$\Omega_n = \begin{vmatrix} \alpha_1 & -M_1 & -M_1 & . & . & . & -M_1 \\ -m_2 & \alpha_2 & -M_2 & . & . & . & -M_2 \\ -m_3 & -m_3 & \alpha_3 & . & . & . & -M_3 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ -m_n & -m_n & -m_n & . & . & . & \alpha_n \end{vmatrix}. \tag{17}$$

If we subtract here the last column from each of the preceding ones, we obtain

$$\begin{vmatrix} M_1+\alpha_1 & 0 & 0 & . & . & . & 0 & -M_1 \\ M_2-m_2 & M_2+\alpha_2 & 0 & . & . & . & 0 & -M_2 \\ M_3-m_3 & M_3-m_3 & M_3+\alpha_3 & . & . & . & 0 & -M_3 \\ . & . & . & & & & . & . \\ . & . & . & & & & . & . \\ . & . & . & & & & . & . \\ M_{n-1}-m_{n-1} & M_{n-1}-m_{n-1} & M_{n-1}-m_{n-1} & . & . & . & M_{n-1}-m_{n-1} & -M_{n-1} \\ -(\alpha_n+m_n) & -(\alpha_n+m_n) & -(\alpha_n+m_n) & . & . & . & -(\alpha_n+m_n) & \alpha_n \end{vmatrix}.$$

Here the subdeterminant corresponding to the last element of the last row is obviously $\prod_{\nu=1}^{n-1} (M_\nu+\alpha_\nu)$, so that Ω_n is the sum of $\alpha_n \prod_{\nu=1}^{n-1} (M_\nu+\alpha_\nu)$, and

$$-(\alpha_n+m_n) \begin{vmatrix} M_1+\alpha_1 & 0 & 0 & . & . & . & 0 & -M_1 \\ M_2-m_2 & M_2+\alpha_2 & 0 & . & . & . & 0 & -M_2 \\ M_3-m_3 & M_3-m_3 & M_3+\alpha_3 & . & . & . & 0 & -M_3 \\ . & . & . & & & & . & . \\ . & . & . & & & & . & . \\ . & . & . & & & & . & . \\ M_{n-1}-m_{n-1} & M_{n-1}-m_{n-1} & M_{n-1}-m_{n-1} & . & . & . & M_{n-1}+\alpha_{n-1} & -M_{n-1} \\ 1 & 1 & 1 & . & . & . & 1 & 0 \end{vmatrix}.$$

This last determinant becomes T_n^* if we put

$$\delta_\nu = M_\nu + \alpha_\nu, \quad \gamma_\nu = M_\nu - m_\nu, \quad \beta_\nu = -M_\nu \qquad (\nu = 1, \ldots, n-1),$$

and has therefore by (16) the value

$$-\sum_{\kappa=1}^{n-1} (-M_\kappa) \prod_{\lambda=1}^{\kappa-1} (M_\lambda + \alpha_\lambda) \prod_{\lambda=\kappa+1}^{n-1} (m_\lambda + \alpha_\lambda).$$

We obtain therefore for Ω_n the expression

$$\Omega_n = \alpha_n \prod_{\nu=1}^{n-1} (M_\nu + \alpha_\nu) - \sum_{\kappa=1}^{n-1} M_\kappa \prod_{\lambda=1}^{\kappa-1} (M_\lambda + \alpha_\lambda) \prod_{\lambda=\kappa+1}^{n} (m_\lambda + \alpha_\lambda). \tag{18}$$

The determinant Ω_n can be also written in the form

$$\Omega_n = \begin{vmatrix}
\alpha_n & -m_n & -m_n & \ldots & -m_n & -m_n \\
-M_{n-1} & \alpha_{n-1} & -m_{n-1} & \ldots & -m_{n-1} & -m_{n-1} \\
-M_{n-2} & -M_{n-2} & \alpha_{n-2} & \ldots & -m_{n-2} & -m_{n-2} \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
-M_2 & -M_2 & -M_2 & \ldots & \alpha_2 & -m_2 \\
-M_1 & -M_1 & -M_1 & \ldots & -M_1 & \alpha_1
\end{vmatrix}, \tag{19}$$

and we obtain therefore from (18)

$$\Omega_n = \alpha_1 \prod_{\nu=2}^{n} (m_\nu + \alpha_\nu) - \sum_{\kappa=2}^{n} m_\kappa \prod_{\mu=\kappa+1}^{n} (m_\mu + \alpha_\mu) \prod_{\mu=1}^{\kappa-1} (M_\mu + \alpha_\mu). \tag{20}$$

3. The Matrix Δ_n and Its Inverse

We consider now the matrix

$$\Delta_n = \begin{pmatrix}
1 & -M & -M & \ldots & -M \\
-m & 1 & -M & \ldots & -M \\
\cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot \\
-m & -m & -m & \ldots & 1
\end{pmatrix} \tag{21}$$

Its determinant is obtained from Ω_n in putting in (20)

$$\alpha_1 = \alpha_2 = \ldots \alpha_n = 1; \quad m_2 = m_3 = \ldots = m_n = m; \quad M_1 = M_2 = \ldots = M_{n-1} = M.$$

We obtain

$$|\Delta_n| = (m+1)^{n-1} - m \sum_{\kappa=2}^{n} (m+1)^{n-\kappa} (M+1)^{\kappa-1},$$

and this becomes, if $M \neq m$,

$$|\Delta_n| = (m+1)^{n-1} - m(M+1) \frac{(M+1)^{n-1} - (m+1)^{n-1}}{M-m}$$

$$= \frac{1}{M-m} [(m+1)^{n-1}(M-m+m(M+1)) - m(M+1)^n],$$

$$|\Delta_n| = \frac{1}{M-m} [M(m+1)^n - m(M+1)^n] \qquad (m \neq M),^6 \tag{22}$$

while for $m = M$ we obtain from (22) in letting $M \to m$:

$$|\Delta_n| = [1 - (n-1)m](1+m)^{n-1} \qquad (m = M). \tag{23}$$

In particular, if $0 < m < M$, a necessary and sufficient condition for $|\Delta_n| > 0$ is

$$\frac{m}{(m+1)^n} < \frac{M}{(M+1)^n}. \tag{24}$$

We assume now in particular

$$0 < m < M. \tag{25}$$

If we introduce the abbreviations

$$M_n = \frac{(1+M)^n - nM - 1}{M}, \quad m_n = \frac{(1+m)^n - nm - 1}{m}, \tag{26}$$

we can write (22) in the form

$$(M-m)|\Delta_n| = M(1+nm+mm_n) - m(1+nM+MM_n) = M - m - mM(M_n - m_n),$$

and therefore, if we put

$$\delta = 1 - |\Delta_n|, \tag{27}$$

$$\frac{\delta}{mM} = \frac{M_n - m_n}{M - m}. \tag{28}$$

It follows from (25) for $n > 2$

$$\frac{M_n - m_n}{M - m} = \sum_{\nu=2}^{n} \binom{n}{\nu} \frac{M^{\nu-1} - m^{\nu-1}}{M - m} > \sum_{\nu=2}^{n} \binom{n}{\nu} M^{\nu-2} = \frac{(1+M)^n - 1 - nM}{M^2} = \frac{M_n}{M},$$

so that from (28) we have

$$\frac{M_n}{1 - \frac{m}{M}} > \frac{\delta}{m} > M_n. \tag{29}$$

It follows in particular that if (24) holds, then $\delta > 0$, $0 < |\Delta_n| < 1$.

In solving (29) with respect to m and δ, we obtain

$$\frac{\delta}{M_n + \frac{\delta}{M}} < m < \frac{\delta}{M_n}, \quad m = \frac{\delta}{M_n + \theta \frac{\delta}{M}}, \quad 0 < \theta < 1, \tag{30}$$

$$M_n m < \delta < \frac{M_n m}{1 - \frac{m}{M}}, \quad \delta = \frac{M_n m}{1 - \theta \frac{m}{M}}, \quad 0 < \theta_1 < 1. \tag{30°}$$

---

^{6} This formula can be also obtained from the formulas given in the proof of theorem III in [3a, p. 113].

In order to discuss the meaning of the condition (24), consider the curve

$$y = f(x) = \frac{x}{(1+x)^n}; \tag{31}$$

we have

$$f'(x) = \frac{1-(n-1)x}{(1+x)^{n+1}}, \quad f''(x) = n\,\frac{(n-1)x-2}{(1+x)^{n+2}}.$$

For $x > 0$ the curve (31) has one maximum at $x = 1/(n-1)$ and an inflection point at $x = 2/(n-1)$.

In figure 1 the curve (31) is drawn (computed for $n=5$). The portion of this curve from the point 0 to the highest point T will be denoted as the *ascending branch* and the portion between T and $x = \infty$ as the *descending branch*.
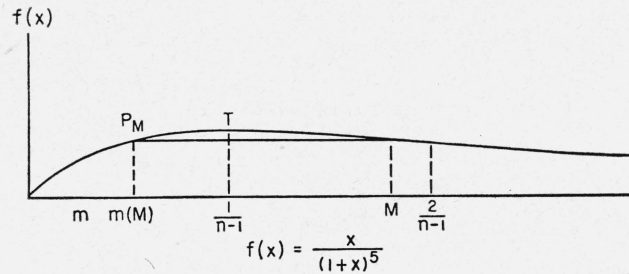

$$f(x) = \frac{x}{(1+x)^5}$$

FIGURE 1.

If we assume that in (24) m is less than M, we see that either

$$0 < m < \frac{1}{n-1}, \quad M > \frac{1}{n-1} \tag{32}$$

or

$$m < M \leqq \frac{1}{n-1}. \tag{32°}$$

In order to find for a given $M > 1/(n-1)$ the range of values of m satisfying (24) we find on the curve the second point P_M with the same ordinate as the point $x = M$; then if the abcissa $m(M)$ corresponds to P_M, we have $0 < m < m(M)$.

Finally we prove from (24) and $m < M$ that it always follows that

$$mM^{n-1} < 1. \tag{33}$$

Indeed, in proving (33) we can obviously assume that M is greater than 1 and therefore on the descending branch in the diagram. But then we have

$$\frac{1}{1+\dfrac{1}{M^{n-1}}} = \frac{M^{n-1}}{1+M^{n-1}} > \frac{M}{1+M},$$

and therefore in raising it to the nth power

$$\frac{\dfrac{1}{M^{n-1}}}{\left(1+\dfrac{1}{M^{n-1}}\right)^n} > \frac{M}{(1+M)^n}.$$

We see that $1/M^{n-1}$ must lie, in figure 1, between M and $m(M)$, and it follows that $1/M^{n-1} > m$, that is (33).

We are now going to determine the inverse matrix

$$\Delta_n^{-1} = (\alpha_{\mu\nu}) \qquad (\mu, \nu = 1, \ldots, n) \tag{34}$$

of Δ_n, if its determinant (22) resp. (23) is not equal to 0. It is sufficient for this purpose to solve the corresponding linear system

$$x_\mu - m \sum_{\nu=1}^{\mu-1} x_\nu - M \sum_{\nu=\mu+1}^{n} x_\nu = z_\mu \qquad (\mu = 1, \ldots, n) \tag{35}$$

for indeterminate z_μ. Put

$$\left.\begin{array}{ll} x_1 + \ldots + x_n = s, & \dfrac{m-M}{M+1} = \sigma, \\[2mm] \dfrac{M}{1+M} = \lambda, & 1 + \sigma = \dfrac{1+m}{1+M} = q, \\[2mm] \dfrac{z_\nu}{1+M} = \alpha_\nu, & \alpha_\nu + \lambda s = y_\nu \end{array}\right\}. \tag{36}$$

Then the system (35) is equivalent with

$$(1+M)x_\mu = (m-M)(x_1 + \ldots + x_{\mu-1}) + Ms + z_\mu \qquad (\mu = 1, \ldots, n),$$

$$x_\mu = \sigma(x_1 + \ldots + x_{\mu-1}) + y_\mu \qquad (\mu = 1, \ldots, n),$$

or in putting

$$S_\nu = x_1 + \ldots + x_\nu \quad (\nu = 1, \ldots, n), \qquad S_0 = 0:$$

$$S_\mu - S_{\mu-1} = \sigma S_{\mu-1} + y_\mu, \qquad S_\mu = (1+\sigma)S_{\mu-1} + y_\mu = q S_{\mu-1} + y_\mu,$$

$$S_\mu q^{-\mu} = S_{\mu-1} q^{-\mu+1} + y_\mu q^{-\mu} \qquad (\mu = 1, \ldots, n),$$

$$S_\mu q^{-\mu} = \sum_{\nu=1}^{\mu} y_\nu q^{-\nu},$$

$$S_\mu = \sum_{\nu=1}^{\mu} y_\nu q^{\mu-\nu} \qquad (\mu = 1, \ldots, n), \tag{37}$$

$$x_\mu = \sum_{\nu=1}^{\mu-1} y_\nu (q-1) q^{\mu-\nu-1} + y_\mu \qquad (\mu = 1, \ldots, n). \tag{38}$$

From (37) for $\mu = n$ we have by (36)

$$S_n = s = \sum_{\nu=1}^{n} y_\nu q^{n-\nu} = \lambda s \sum_{\nu=1}^{n} q^{n-\nu} + \sum_{\nu=1}^{n} \alpha_\nu q^{n-\nu},$$

$$s = \lambda s \frac{q^n - 1}{q-1} + \sum_{\nu=1}^{n} \alpha_\nu q^{n-\nu}, \tag{39}$$

where for $m = M$, $q = 1$ the coefficient of λs is to be replaced by n. Since for $m \neq M$

$$1 - \lambda \frac{q^n - 1}{q-1} = \frac{m - Mq^n}{m-M} = \frac{Mq^n - m}{M-m},$$

we have from (39) and (22)

$$s = Q \sum_{\nu=1}^{n} \alpha_\nu q^{n-\nu} \qquad (m \neq M), \tag{40}$$

$$Q = \frac{M-m}{Mq^n - m} = \frac{(M+1)^n}{|\Delta_n|} \qquad (m \neq M), \tag{41}$$

while for $m=M$ we obtain again (40) in defining Q by

$$Q=\frac{m+1}{1-(n-1)m}=\frac{(m+1)^n}{|\Delta_n|}\qquad(m=M).\tag{41°}$$

If we now replace in (38) y_ν by $\lambda s+\alpha_\nu$ and use for s the expression from (40),

$$x_\mu=(q-1)\sum_{\nu=1}^{\mu-1}\alpha_\nu q^{\mu-\nu-1}+\alpha_\mu+\lambda s\left[(q-1)\sum_{\nu=1}^{\mu-1}q^{\mu-\nu-1}+1\right].$$

But here the expression within the brackets is $q^{\mu-1}$, and we obtain therefore

$$x_\mu=\sum_{\nu=1}^{\mu-1}\alpha_\nu q^{\mu-\nu-1}(q-1)+\alpha_\mu+\lambda Q q^{\mu-1}\sum_{\nu=1}^{n}\alpha_\nu q^{n-\nu}\qquad(\mu=1,\ldots,n)$$

and finally in introducing α_μ from (36)

$$x_\mu=\sum_{\nu=1}^{\mu-1}z_\nu q^{\mu-\nu-1}\frac{q-1}{M+1}+\frac{z_\mu}{M+1}+\frac{\lambda Q q^{\mu-1}}{M+1}\sum_{\nu=1}^{n}z_\nu q^{n-\nu}\qquad(\mu=1,\ldots,n).\tag{42}$$

Introduce the two following triangular matrices:

$$U_n(q)=(u_{\mu\nu})=\begin{pmatrix}0 & 0 & . & . & . & . & . & . & 0\\ 1 & 0 & . & . & . & . & . & . & .\\ q & 1 & 0 & & & & & & .\\ q^2 & q & . & . & & & & & .\\ . & . & . & . & & & & & .\\ . & . & . & . & . & & & & .\\ . & . & . & . & . & . & & & .\\ . & . & . & . & . & . & . & & .\\ . & . & . & . & q^2 & q & 1 & 0\end{pmatrix},\tag{43}$$

$$u_{\mu\nu}=\begin{cases}q^{\mu-\nu-1} & (\mu>\nu)\\ 0 & (\mu\leqq\nu),\end{cases}\tag{44}$$

$$V_n(q)=(v_{\mu\nu})=\begin{pmatrix}0 & q^{-1} & q^{-2} & q^{-3} & \ldots & q^{-n+1}\\ 0 & 0 & q^{-1} & q^{-2} & \ldots & q^{-n+2}\\ . & . & . & . & . & \ldots . .\\ . & . & . & . & . & \ldots . .\\ . & . & . & . & . & \ldots . .\\ 0 & 0 & 0 & 0 & \ldots & 0\end{pmatrix},\tag{45}$$

$$v_{\mu\nu}=\begin{cases}0 & (\mu\geqq\nu)\\ q^{\mu-\nu} & (\mu<\nu).\end{cases}\tag{46}$$

We see then that in (42) the matrix corresponding to the first right-hand sum is

$$\frac{q-1}{M+1}\,U_n=\frac{m-M}{(M+1)^2}\,U_n,$$

whereas to the second right-hand sum in (42) corresponds the matrix

$$\frac{\lambda Q}{M+1} \, q^{n-1}(q\,U_n + E_n + V_n),$$

where E_n is the unit matrix. We obtain

$$\Delta_n^{-1} = \left(\frac{m-M}{(M+1)^2} + \frac{MQq^n}{(M+1)^2}\right)U_n + \left(\frac{1}{M+1} + \frac{MQq^{n-1}}{(M+1)^2}\right)E_n + \frac{MQq^{n-1}}{(M+1)^2}\,V_n.$$

The coefficient of U_n is here equal to $(mQ)/(M+1)^2$, both in the case (41) and (41°), and we have finally

$$\Delta_n^{-1} = \frac{mQ}{(M+1)^2}\,U_n + \left(\frac{1}{M+1} + \frac{MQq^{n-1}}{(M+1)^2}\right)E_n + \frac{MQq^{n-1}}{(M+1)^2}\,V_n. \tag{47}$$

We now denote by s_μ the sum of the elements in the μth row of the matrix (47) and by t_μ the corresponding sum in the μth column. If we first assume $m \neq M$, for the matrices U_n and V_n, the μth row sums are, respectively, equal to

$$\frac{q^{\mu-1}-1}{q-1}; \qquad \frac{1-q^{\mu-n}}{q-1},$$

and the νth column sums are

$$\frac{q^{n-\nu}-1}{q-1}; \qquad \frac{1-q^{1-\nu}}{q-1}.$$

We obtain now from (47) by (36)

$$(M+1)^2 s_\mu = mQ\,\frac{q^{\mu-1}-1}{q-1} + M+1 + MQq^{n-1} + MQ\,\frac{q^{n-1}-q^{\mu-1}}{q-1}$$

$$= \frac{Q}{\sigma}\,[mq^{\mu-1} - m + Mq^n - Mq^{n-1} + Mq^{n-1} - Mq^{\mu-1}] + M+1$$

$$= \frac{M+1}{m-M}\,Q\,[Mq^n - m + (m-M)q^{\mu-1}] + M+1,$$

and this is by (41) equal to $(M+1)Qq^{\mu-1}$. We have finally

$$s_\mu = \frac{Q}{M+1}\,q^{\mu-1} = \frac{(M+1)^{n-1}}{|\Delta_n|}\,q^{\mu-1}, \tag{48}$$

and this remains true also for $m = M$, as is immediately seen.

Similarly, we have

$$(M+1)^2 t_\nu = mQ\,\frac{q^{n-\nu}-1}{q-1} + M+1 + MQq^{n-1} + MQ\,\frac{q^{n-1}-q^{n-\nu}}{q}$$

$$= \frac{Q}{\sigma}\,[mq^{n-\nu} - m + Mq^n - Mq^{n-1} + Mq^{n-1} - Mq^{n-\nu}] + M+1$$

$$= \frac{M+1}{m-M}\,Q\,[Mq^n - m + (m-M)q^{n-\nu}] + M+1 = (M+1)\,Qq^{n-\nu},$$

and we have

$$t_\nu = \frac{Q}{M+1}\,q^{n-\nu} = \frac{(M+1)^{n-1}}{|\Delta_n|}\,q^{n-\nu}, \tag{49}$$

relation which is also immediately verified for the case $m = M$.

4. Bounds for the Determinants, depending on Ω_n and Δ_n

A determinant

$$M=\begin{vmatrix} \alpha_1 & -m_{12} & \ldots & -m_{1n} \\ -m_{21} & \alpha_2 & \ldots & -m_{2n} \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ -m_{n1} & -m_{n2} & \ldots & \alpha_n \end{vmatrix} \tag{50}$$

will be called an *M-determinant* if all diagonal elements α_ν are positive, all elements off the main diagonal $-m_{\mu\nu}$ $(\mu \neq \nu)$ are not positive and the determinant M, as well as all principal (coaxial) minors of M of all degrees, are positive.

In what follows we will have to use a theorem given by the author in 1937 [1], and which will be formulated as the following:

LEMMA I. *If we have for the M-determinant (50) and a determinant $H = |h_{\mu\nu}|$ $(\mu, \nu = 1, \ldots, n)$ of order n, the inequalities*

$$|h_{\nu\nu}| \geqq \alpha_\nu, \qquad |h_{\mu\nu}| \leqq m_{\mu\nu} \qquad (\mu \neq \nu; \mu, \nu = 1, \ldots, n), \tag{51}$$

then $H \neq 0$, and we have

$$|H| \geqq M. \tag{52}$$

If (H), (M), respectively, denote the matrices of the determinants H and M, the inverse matrix of (H) is majorized by the inverse matrix of (M)

$$(H)^{-1} \ll (M)^{-1}. \tag{52°}$$

Suppose now that in the determinant Ω_n given by (17) all α_ν are positive and m_ν, M_ν nonnegative. Then it follows from (18) that Ω_n is a monotonically decreasing function of all m_ν and from (20) that M is also monotonically decreasing in all M_ν. Suppose now that for a certain set of values of α_ν, m_ν, M_ν the determinant $\Omega_n \neq 0$. Then replace the m_ν and M_ν corresponding to certain rows $(\nu_1, \ldots, \nu_\kappa)$ by zeros; the determinant Ω_n cannot decrease and remains therefore positive; but then Ω_n becomes equal to the product of $\alpha_{\nu_1}\alpha_{\nu_2} \ldots \alpha_{\nu_\kappa}$ with the principal minor complementary to the set of indices $\nu_1, \ldots, \nu_\kappa$. Therefore, all principal minors of Ω_n are positive. We see that Ω_n *is an M-determinant, if the conditions*

$$\alpha_\nu > 0, \quad m_\nu \geqq 0, \quad M_\nu \geqq 0 \qquad (\nu = 1, \ldots, n); \tag{53}$$

are satisfied and $\Omega_n \neq 0$. But now we can easily deduce the following theorem.

A. *Consider the set of all determinants $A = |a_{\mu\nu}|$ satisfying the conditions*

$$|a_{\mu\nu}| \leqq m_\mu \quad (\nu < \mu), \qquad |a_{\mu\nu}| \leqq M_\mu \quad (\nu > \mu), \qquad |a_{\mu\mu}| \geqq \alpha_\mu \quad (\mu, \nu = 1, \ldots, n), \tag{54}$$

where α_μ are n given positive constants and m_μ, M_μ are $2n-2$ given nonnegative constants. Then in order that no determinant A of this set vanishes, it is necessary and sufficient that $\Omega_n > 0$. If this condition is satisfied, we have

$$|A| \geqq \Omega_n. \tag{55}$$

PROOF: Since in the case $\Omega_n > 0$, Ω_n is an M-determinant, the sufficiency of our condition follows immediately from the inequality (52) mentioned above.

If $\Omega_n = 0$, we can take obviously $a_{\mu\mu} = \alpha_\mu$ and $a_{\mu\nu} = -m_\mu$ or $a_{\mu\nu} = -M_\mu$ according as $\nu < \mu$ or $\nu > \mu$ and obtain a vanishing determinant A satisfying the condition (54).

Suppose now $\Omega_n < 0$, then it follows from the form (17) of Ω_n that Ω_n becomes positive if all m_μ are replaced by zeros. There exists, therefore, such a positive $t < 1$ that Ω_n vanishes if all m_μ in (17) are replaced by tm_μ, but then we obtain a vanishing determinant A of our set in taking $a_{\mu\mu} = \alpha_\mu$ and $a_{\mu\nu} = -tm_\mu$ or $a_{\mu\nu} = -M_\mu$, according as $\nu < \mu$ or $\nu > \mu$. The theorem A is proved.

In specializing the matrix of Ω_n to Δ_n and in assuming that in particular $0 < m < M$, we obtain immediately from (22) and the theorem A:

B. *Consider for two positive constants m, M, $m \leqq M$, the set of all determinants $A = |a_{\mu\nu}|$ of the nth order for which*

$$|a_{\mu\mu}| \geqq 1 \quad (\mu = 1, \ldots, n), \quad |a_{\mu\nu}| \leqq m \quad (\nu < \mu), \quad |a_{\mu\nu}| \leqq M \quad (\nu > \mu), \qquad (56)$$

then, in order that no determinant A of this set vanish, it is necessary and sufficient that for $m < M$ the inequality

$$\frac{m}{(1+m)^n} < \frac{M}{(1+M)^n} \qquad (m < M), \qquad (57)$$

holds, and if this inequality is satisfied, we have for each determinant A of the set

$$|A| \geqq |\Delta_n| = \frac{(1+m)^n (1+M)^n}{M-m} \left[\frac{M}{(1+M)^n} - \frac{m}{(1+m)^n} \right]. \qquad (58)$$

If $m = M$, the condition becomes

$$m < \frac{1}{n-1} \qquad (M = m) \qquad (57°)$$

and we have, if (57°) is satisfied,

$$|A| \geqq |\Delta_n| = [1 - (n-1)m](1+m)^{n-1} \qquad (m = M). \qquad (58°)$$

From theorem B we can deduce the following theorem.

C. *Let $A = (a_{\mu\nu})$ $(\mu, \nu = 1, \ldots, n)$ be a matrix satisfying the conditions*

$$|a_{\mu\nu}| \leqq m \quad (\nu < \mu), \qquad |a_{\mu\nu}| \leqq M \quad (\nu > \mu), \qquad (59)$$

where m, M are two constants with $0 < m < M$. Put

$$\delta(m, M) = \frac{Mm^{1/n} - mM^{1/n}}{M^{1/n} - m^{1/n}}. \qquad (60)$$

Then all fundamental roots of the matrix A are contained in the set of the n closed circles described around the elements $a_{\mu\mu}$ with the radius $\delta(m, M)$. The value (60) of $\delta(m, M)$ cannot be improved if $A = \Delta_n$.[7]

PROOF: Let λ be a fundamental root of A so that the matrix $\lambda E - A$ is singular. Put $\min_\mu |\lambda - a_{\mu\mu}| = \alpha$; we then have to prove that $\alpha \leqq \delta(m, M)$. If $\alpha = 0$, there is nothing to prove. Suppose $\alpha > 0$, and consider the matrix

$$\frac{\lambda E - A}{\alpha} = (b_{\mu\nu}).$$

For this matrix we have

$$|b_{\mu\mu}| \geqq 1, \quad |b_{\mu\nu}| \leqq \frac{m}{\alpha} \quad (\nu < \mu), \qquad |b_{\mu\nu}| \leqq \frac{M}{\alpha} \quad (\nu > \mu).$$

---

[7] For $M = m$, $\delta(m, M)$ becomes $(n-1)m$, but in this case our result is contained in Gerschgorin's theorem, cf. [4].

From the theorem C the result given by Stein and Rosenberg in [3a] as theorem III follows immediately.

Therefore, we have by theorem B

$$\frac{\dfrac{m}{\alpha}}{\left(1+\dfrac{m}{\alpha}\right)^n} \geqq \frac{\dfrac{M}{\alpha}}{\left(1+\dfrac{M}{\alpha}\right)^n},$$

and therefore,

$$\frac{1+\dfrac{m}{\alpha}}{m^{1/n}} \leqq \frac{1+\dfrac{M}{\alpha}}{M^{1/n}},$$

$$\alpha\left(\frac{1}{m^{1/n}}-\frac{1}{M^{1/n}}\right) \leqq \frac{M}{M^{1/n}}-\frac{m}{m^{1/n}},$$

$$\alpha \leqq \frac{M^{1-\frac{1}{n}}-m^{1-\frac{1}{n}}}{m^{-1/n}-M^{-1/n}} = \delta(m, M).$$

5. Bounds of the Matrix Δ_n^{-1}

For an n dimensional vector $\xi=(x_1, \ldots, x_n)$ the Hölder norm corresponding to the exponent $p \geqq 1$ is given by

$$|\xi|_p = \sqrt[p]{|x_1|^p + \ldots + |x_n|^p} \qquad (p \geqq 1). \tag{61}$$

We will only use the three cases corresponding to $p=1, 2, \infty$:

$$|\xi|_p = |x_1| + \ldots + |x_n|, \quad |\xi|_\infty = \max_\nu |x_\nu|, \quad |\xi|_2 = \sqrt{|x_1|^2 + \ldots + |x_n|^2}. \tag{62}$$

We have among these three norms the following inequalities:

$$|\xi|_\infty \leqq |\xi|_1 \leqq n|\xi|_\infty, \tag{63,a}$$

$$|\xi|_\infty \leqq |\xi|_2 \leqq \sqrt{n}|\xi|_\infty, \tag{63,b}$$

$$|\xi|_1 \leqq \sqrt{n}|\xi|_2 \leqq \sqrt{n}|\xi|_1, \tag{63,c}$$

which are immediately verified. The left-hand inequality (63, c) implies the well-known inequality between the arithmetical mean and the arithmetical mean of the squares. If $A=(a_{\mu\nu})$ is an $n \times n$ matrix, we define its norm corresponding to the exponent p $(1 \leqq p \leqq \infty)$ by

$$|A|_p = \max_{\xi \neq 0} \frac{|A\xi|_p}{|\xi|_p} \tag{64}$$

and denote it by $|A|_p$. We will use here too only the cases $p=1, 2, \infty$.

In applying to the definition (64) the formulas (63, b), and (63, c) we obtain immediately

$$\frac{1}{\sqrt{n}}|A|_1 \leqq |A|_2 \leqq \sqrt{n}\,|A|_1, \tag{65,a}$$

$$\frac{1}{\sqrt{n}}|A|_\infty \leqq |A|_2 \leqq \sqrt{n}\,|A|_\infty. \tag{65,b}$$

For $p=1, \infty$ the expressions of $|A|_1, |A|_\infty$ are easy to write down; we have, as is well known and very easy to prove,

$$|A|_1 = \max_\nu \sum_\mu |a_{\mu\nu}| \tag{66,a}$$

$$|A|_\infty = \max_\mu \sum_\nu |a_{\mu\nu}|. \tag{66,b}$$

As to $|A|_2$, its expression is irrational; $|A|_2$ is the square root of the maximum fundamental root of the symmetric and nonnegative matrix AA^*. Since the direct computation of $|A|_2$ is difficult in most cases, we prove the following estimate for $|A|_2$:

LEMMA II. *We have for any matrix A*

$$\frac{1}{n}\max\left(|A|_1, |A|_\infty\right) \leqq |A|_2 \leqq \sqrt{|A|_1 |A|_\infty}, \tag{67}$$

The first part of (67) follows from (65,a) and (65,b). To prove the second part, we introduce the notations

$$s_\mu = \sum_\nu |a_{\mu\nu}|, \quad t_\nu = \sum_\mu |a_{\mu\nu}| \quad (\mu, \nu = 1, \ldots, n).$$

The sum of the moduli of all elements in the μth row of AA^* can be estimated as follows

$$\left|\sum_{\nu, \kappa} a_{\mu\kappa} a_{\nu\kappa}\right| \leqq \sum_\kappa \sum_\nu |a_{\mu\kappa} a_{\nu\kappa}| = \sum_\kappa |a_{\mu\kappa}| t_\kappa \leqq s_\mu |A|_1 \leqq |A|_\infty |A|_1.$$

Equation (67) follows now from the theorem of Frobenius that the modulus of each fundamental root of a square matrix does not exceed the greatest sum of the moduli of the elements of this matrix in different rows.

If we now apply these results to the matrix Δ_n^{-1} discussed in the section 3, we obtain from (48), (49), (66,a) and (66,b)

$$|\Delta_n^{-1}|_1 = |\Delta_n^{-1}|_\infty = \frac{Q}{M+1} = \frac{(M+1)^{n-1}}{|\Delta_n|}, \tag{68}$$

and from (67) and (65,a)

$$\frac{Q}{\sqrt{n}(M+1)} \leqq |\Delta_n^{-1}|_2 \leqq \frac{Q}{M+1} = \frac{(M+1)^{n-1}}{|\Delta_n|}. \tag{69}$$

In combining these inequalities with the results given in section 4, we obtain the following theorem.

D. *Let $A = (a_{\mu\nu})$ be a square matrix of order n satisfying conditions (56), and let (57) be satisfied. Then we have*

$$|A^{-1}|_p \leqq \frac{(M+1)^{n-1}}{|\Delta_n|} \qquad (p=1, 2, \infty);$$

and therefore for any vector ξ

$$|A\xi|_p \geqq \frac{|\Delta_n|}{(M+1)^{n-1}} |\xi|_p \qquad (p=1, 2, \infty). \tag{70}$$

PROOF: It follows from (52°) of lemma I at once in virtue of (66,a) and (66,b) that

$$|A^{-1}|_p \leqq |\Delta_n^{-1}|_p = \frac{Q}{M+1} \qquad (p=1, \infty);$$

further, the matrix $A^{-1}(A^{-1})^*$ is majorized by $\Delta_n^{-1}(\Delta_n^{-1})^*$. By a well-known theorem of Frobenius, therefore, the maximum modulus of a fundamental root of $A^{-1}(A^{-1})^*$ is majorized by that of $\Delta_n^{-1}(\Delta_n^{-1})^*$, and so we also have

$$|A^{-1}|_2 \leqq |\Delta_n^{-1}|_2 \leqq \frac{Q}{M+1}.$$

By definition (64), therefore, in putting $A\xi = \eta$, we have

$$|A^{-1}\eta|_p \leqq \frac{Q}{M+1} |\eta|_p,$$

and this is equivalent to (70).

333

6. Bounds of the Matrix $\Delta_n^{-1}-\Delta_n^{(0)\,-1}$

We denote by $\Delta_n^{(0)}$ the matrix obtained from (21) in replacing m by zero:

$$\Delta^{(0)}=\begin{pmatrix} 1 & -M & -M & \ldots & -M \\ 0 & 1 & -M & \ldots & -M \\ \cdot & \cdot & \cdot & \cdot\cdot\cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot\cdot\cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot\cdot\cdot & \cdot \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}. \tag{71}$$

We will prove the following theorem:

E. *If $0<m<M$ and (57) is satisfied, then the matrix $\Delta_n^{-1}-\Delta_n^{(0)\,-1}$ is a nonnegative matrix and we have*

$$|\Delta_n^{-1}-\Delta_n^{(0)\,-1}|_p=\max_{1\leqq\nu\leqq n}(1+M)^{\nu-1}\frac{(1+m)^{n-\nu}-|\Delta_n|}{|\Delta_n|} \qquad (p=1,\infty). \tag{72}$$

Proof: Since $\Delta_n^{(0)}$ satisfies the inequalities (51), it follows from (52°) that $\Delta_n^{-1}-\Delta_n^{(0)\,-1}$ is nonnegative. Denote by $s_\mu^{(0)}$ and $t_\nu^{(0)}$ the sums of the elements in the μth row and in the νth column of $\Delta_n^{(0)\,-1}$, and by $\bar{s}_\mu$ and $\bar{t}_\nu$ the corresponding expressions for the matrix $\Delta_n^{-1}-\Delta_n^{(0)\,-1}$. We have

$$\bar{s}_\mu=s_\mu-s_\mu^{(0)}, \qquad \bar{t}_\nu=t_\nu-t_\nu^{(0)} \qquad (\mu,\nu=1,\ldots,n). \tag{73}$$

It follows then from the formulas (48) and (49) that $\bar{s}_\mu$ runs for $\mu=n, n-1, \ldots, 1$ through the same set of values as $\bar{t}_\nu$ for $\nu=1, \ldots, n$.

Formula (49) can be written in the form

$$t_\nu=\frac{q^{n-\nu}(M+1)^{n-1}}{|\Delta_n|}=\frac{(1+m)^{n-\nu}(1+M)^{\nu-1}}{|\Delta_n|}. \tag{74}$$

Since Δ_n becomes 1 for $m=0$, we have (72) and E is proved.

Discussion of $|\Delta_n^{-1}-\Delta_n^{(0)\,-1}|_p$. We prove first that $t_\nu-t_\nu^{(0)}$ increases with ν as long as the condition

$$1-\frac{m}{M}\geqq|\Delta_n| \tag{75}$$

is satisfied. Indeed, put

$$k_\nu\equiv\frac{t_{\nu+1}-t_{\nu+1}^{(0)}}{t_\nu-t_\nu^{(0)}}=(M+1)\frac{(1+m)^{n-\nu-1}-|\Delta_n|}{(1+m)^{n-\nu}-|\Delta_n|}. \tag{76}$$

In solving the three inequalities

$$k_\nu\gtreqless 1 \tag{77}$$

with respect to $|\Delta_n|$, we obtain correspondingly

$$|\Delta_n|\lesseqgtr\sigma_\nu\equiv\left(1-\frac{m}{M}\right)(1+m)^{n-\nu-1}. \tag{78}$$

The inequalities $k_\nu\geqq 1$ $(\nu=1,\ldots,n)$ are obviously satisfied in virtue of (75) as long as

334

$\nu \leqq n-1$. We have, therefore, in this case, in using (72)

$$\text{Max}_{\nu}\ (t_\nu - t_\nu^{(0)}) = (1+M)^{n-1}\frac{1-|\Delta_n|}{|\Delta_n|} = (1+M)^{n-1}\frac{\delta}{1-\delta},$$

$$|\Delta_n^{-1} - \Delta_n^{(0)-1}|_p = (1+M)^{n-1}\frac{1-|\Delta_n|}{|\Delta_n|} = (1+M)^{n-1}\frac{\delta}{1-\delta} \qquad \left(p=1,\infty\ ;\quad |\Delta_n| \leqq 1-\frac{m}{M}\right). \quad (79)$$

Condition (75) can be written by using the notation (27), as $\delta \geqq m/M$, and in virtue of (29) this is certainly satisfied if we have $MM_n \geqq 1$,

$$(M+1)^n - nM - 2 \geqq 0. \tag{80}$$

Condition (80) becomes for $n=3$: $M \geqq 0.5321$. *For $n \geqq 4$, (80) is in any case satisfied, if we have*

$$M \geqq \frac{1.5}{n} \qquad (n \geqq 4). \tag{81}$$

Indeed, if the relation (80) is satisfied for a positive M, it is satisfied for any greater value because the coefficients of all positive powers of M in the left-hand expression are greater than or equal to 0. To prove the sufficiency of (81), it is sufficient to prove that $(1+1.5/n)^n \geqq 3.5$ $(n \geqq 4)$. But here the left-hand expression is monotonically increasing with n, and this inequality follows, therefore, from $(5.5/4)^4 = 3.75 \ldots > 3.5$.

Suppose now that we have

$$1 - \frac{m}{M} \leqq |\Delta_n| \leqq 1, \quad \delta \leqq \frac{m}{M}. \tag{82}$$

Then we have, since $0 < m < 1/(n-1)$,

$$\frac{(1+m)^{n-\nu}-1}{m} \leqq \frac{\left(1+\dfrac{1}{n-1}\right)^{n-\nu}-1}{\dfrac{1}{n-1}} \leqq (n-1)\left(\left(1+\frac{1}{n-1}\right)^{n-1}-1\right) \leqq (e-1)(n-1),$$

$$(1+m)^{n-\nu} = 1 + 1.72\,\theta nm, \quad 0 < \theta < 1$$

$$(1+m)^{n-\nu} - |\Delta_n| = 1.72\,\theta nm + 1 - |\Delta_n|,$$

and therefore by (82) $(1+m)^{n-\nu} - |\Delta_n| \leqq 1.72\theta nm + m/M$.

For $n \geqq 4$ we have in the case (82), since the inequality (81) is not satisfied,

$$\frac{1}{M} \leqq \frac{n}{1.5}$$

and, therefore, finally

$$(1+m)^{n-\nu} - |\Delta_n| \leqq 2.39nm \qquad (\nu = 1, \ldots, n) \tag{83}$$

On the other hand, since $M < 1.5/n$, $(1+M)^{\nu-1} \leqq (1+1.5/n)^n < e^{1.5} = 4.481689$, and, therefore, from (72) and (73)

$$t_\nu - t_\nu^{(0)} \leqq 10.72n\frac{m}{|\Delta_n|} \quad \left(\nu=1,\ldots,n;\quad |\Delta_n| \geqq 1-\frac{m}{M};\quad n>3\right),$$

$$|\Delta_n^{-1} - \Delta_n^{(0)-1}|_p \leqq 10.72n\frac{m}{|\Delta_n|} \quad \left(p=1,\infty\ ;\quad |\Delta_n| \geqq 1-\frac{m}{M};\quad n>3\right). \tag{84}$$

To obtain a lower bound, we take in (74) $\nu=1$. We obtain, since for $n>3$, $0 < |\Delta_n| \leqq 1$,

$$t_1 - t_1^{(0)} = \frac{(1+m)^{n-1}-|\Delta_n|}{|\Delta_n|} \geqq \frac{(n-1)m}{|\Delta_n|} \geqq 0.75n\frac{m}{|\Delta_n|},$$

and, therefore,

$$|\Delta_n^{-1} - \Delta_n^{(0)-1}|_p \geqq 0.75n \frac{m}{|\Delta_n|} \qquad (p = 1, \infty; n \geqq 4). \tag{85}$$

To obtain the exact value of $|\Delta_n^{-1} - \Delta_n^{(0)-1}|_p$ if (75) is not satisfied, we return to the inequalities (78) which are equivalent to (77).

Denote by n_0 the smallest integer between 1 and n, such that we have

$$\sigma_{n_0-1} \geqq |\Delta_n| > \sigma_{n_0} \qquad (1 \leqq n_0 \leqq n). \tag{86}$$

The parts of this inequality implying σ_0 or σ_n must be disregarded, that is to say, this inequality reduces to $|\Delta_n| > \sigma_1$ for $n_0 = 1$ and to $\sigma_{n-1} \geqq |\Delta_n|$ for $n_0 = n$. Then we see at once that $\max (t_\nu - t_\nu^{(0)}) = t_{n_0} - t_0^{(0)}$, and therefore

$$|\Delta_n^{-1} - \Delta_n^{(0)-1}|_p = t_{n_0} - t_0^{(0)} = \frac{(1+M)^{n_0-1}(1+m)^{n-n_0} - |\Delta_n|}{|\Delta_n|} \qquad (p = 1, \infty). \tag{87}$$

For $m = M$ we have $n_0 = 1$, and therefore by (23)

$$|\Delta_n^{-1} - \Delta_n^{(0)-1}|_p = \frac{(n-1)m}{1-(n-1)m} \qquad \left(m = M < \frac{1}{n-1} \right). \tag{87°}$$

7. Bounds of $A^{-1} - A^{(0)-1}$

In order to obtain the theorem corresponding to E for a determinant A satisfying conditions (56), (57) of theorem B, we prove first the following important lemma, which generalizes considerably the relation (52°) of lemma I.

LEMMA III. *Consider n positive numbers $\alpha_1, \ldots, \alpha_n$ and $2n^2 - 2n$ nonnegative numbers $m_{\mu\nu}, \epsilon_{\mu\nu}$ ($\mu \neq \nu$; $\mu, \nu = 1, \ldots, n$), such that the matrices*

$$M = \begin{bmatrix} \alpha_1 & -m_{12}-\epsilon_{12} & \cdots & -m_{1n}-\epsilon_{1n} \\ -m_{21}-\epsilon_{21} & \alpha_2 & \cdots & -m_{2n}-\epsilon_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ -m_{n1}-\epsilon_{n1} & -m_{n2}-\epsilon_{n2} & \cdots & \alpha_n \end{bmatrix}, \tag{88}$$

$$M^{(0)} = \begin{bmatrix} \alpha_1 & -m_{12} & \cdots & -m_{1n} \\ -m_{21} & \alpha_2 & \cdots & -m_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ -m_{n1} & -m_{n2} & \cdots & \alpha_n \end{bmatrix} \tag{89}$$

are M-matrices.

Consider n constants A_μ ($\mu = 1, \ldots, n$), such that

$$|A_\mu| \geqq \alpha_\mu \qquad (\mu = 1, \ldots, n) \tag{90}$$

336

and $2n^2-2n$ constants $a_{\mu\nu}$, $b_{\mu\nu}$ $(\mu\neq\nu;\quad \mu,\nu=1,\ldots,n)$ such that

$$|a_{\mu\nu}|\leqq m_{\mu\nu},\quad |b_{\mu\nu}-a_{\mu\nu}|\leqq \epsilon_{\mu\nu}\quad (\mu\neq\nu;\quad \mu,\nu=1,\ldots,n),\tag{91}$$

and form the two matrices

$$A=\begin{pmatrix} A_1 & b_{12} & \ldots & b_{1n} \\ b_{21} & A_2 & \ldots & b_{2n} \\ \cdot & \cdot & \ldots & \cdot \\ \cdot & \cdot & \ldots & \cdot \\ \cdot & \cdot & \ldots & \cdot \\ b_{n1} & b_{n2} & \ldots & A_n \end{pmatrix},\tag{92}$$

$$A^{(0)}=\begin{pmatrix} A_1 & a_{12} & \ldots & a_{1n} \\ a_{21} & A_2 & \ldots & a_{2n} \\ \cdot & \cdot & \ldots & \cdot \\ \cdot & \cdot & \ldots & \cdot \\ \cdot & \cdot & \ldots & \cdot \\ a_{n1} & a_{n2} & \ldots & A_n \end{pmatrix}.\tag{93}$$

Then we have

$$A^{-1}-A^{(0)-1}\ll M^{-1}-M^{(0)-1}.\tag{94}$$

Proof: We write $M=P-T$, $M^{(0)}=P-T^{(0)}$, where the matrices T, $T^{(0)}$ have zeros in the main diagonal and off the main diagonal, respectively, the nonnegative elements $m_{\mu\nu}+\epsilon_{\mu\nu}$, $m_{\mu\nu}$ and where P is the diagonal matrix

$$\begin{pmatrix} \alpha_1 & & & & \\ & \alpha_2 & & 0 & \\ & & \cdot & & \\ & 0 & & \cdot & \\ & & & & \alpha_n \end{pmatrix}.$$

We can develop the inverse of $P-T=P(E-P^{-1}T)$ (E is the unit matrix) in the following way:

$$(P-T)^{-1}=\sum_{\kappa=0}^{\infty}(P^{-1}T)^{\kappa}P^{-1}.\tag{96}$$

The convergence of this development and the validity of (96) follow easily from the fact that the determinant of the matrix $P-tT$ does not vanish for $|t|\leqq 1$, which follows immediately from lemma I. The corresponding development holds also for $(P-T^{(0)})^{-1}$, and we obtain therefore

$$M^{-1}-M^{(0)-1}=\sum_{\kappa=1}^{\infty}[(P^{-1}T)^{\kappa}-(P^{-1}T^{(0)})^{\kappa}]P^{-1}.\tag{97}$$

The elements of $P^{-1}T$ are here $(m_{\mu\nu}+\epsilon_{\mu\nu})/\alpha_\mu$ or zeros, and those of $P^{-1}T^0$ are $m_{\mu\nu}/\alpha_\mu$ or zeros; therefore all elements of the matrices

$$(P^{-1}T)^\kappa-(P^{-1}T^{(0)})^\kappa \qquad (\kappa=1,2,\ \ldots) \tag{98}$$

are polynomials in $m_{\mu\nu}/\alpha_\mu$ and $\epsilon_{\mu\nu}/\alpha_\mu$ with nonnegative coefficients. Denote by Q the diagonal matrix

$$Q=\begin{bmatrix} A_1 & & & & \\ & A_2 & & 0 & \\ & & \cdot & & \\ & & & \cdot & \\ & 0 & & \cdot & \\ & & & & A_n \end{bmatrix}.$$

nd write $A=Q-S$, $A^{(0)}=Q-S^{(0)}$, where the elements of S are $b_{\mu\nu}$ or zeros, and those of $S^{(0)}$, $a_{\mu\nu}$ or zeros. Since, therefore, we have

$$Q^{-1}S\ll P^{-1}T, \qquad Q^{-1}S^{(0)}\ll P^{-1}T^{(0)},$$

we have

$$A^{-1}=\sum_{\kappa=0}^{\infty}(Q^{-1}S)^\kappa Q^{-1}, \qquad A^{(0)-1}=\sum_{\kappa=0}^{\infty}(Q^{-1}S^{(0)})^\kappa Q^{-1},$$

$$A^{-1}-A^{(0)-1}=\sum_{\kappa=0}^{\infty}[(Q^{-1}S)^\kappa-(Q^{-1}S^{(0)})^\kappa]Q^{-1}. \tag{99}$$

But now the elements of $(Q^{-1}S)^\kappa-(Q^{-1}S^{(0)})^\kappa$ are obtained from those of (98) in substituting there $a_{\mu\nu}/A_\mu$ instead of $m_{\mu\nu}/\alpha_\mu$ and $(b_{\mu\nu}-a_{\mu\nu})/A_\mu$ instead of $\epsilon_{\mu\nu}/\alpha_\mu$. Further, we have by (90) and (91)

$$\left|\frac{a_{\mu\nu}}{A_\mu}\right|\leqq\frac{m_{\mu\nu}}{\alpha_\mu}, \qquad \left|\frac{b_{\mu\nu}-a_{\mu\nu}}{A_\mu}\right|\leqq\frac{\epsilon_{\mu\nu}}{\alpha_\mu},$$

and therefore, since the coefficients in (98) are as already mentioned not negative,

$$(Q^{-1}S)^\kappa-(Q^{-1}S^{(0)})^\kappa\ll(P^{-1}T)^\kappa-(P^{-1}T^{(0)})^\kappa$$

and

$$\sum_{\kappa=1}^{\infty}[(Q^{-1}S)^\kappa-(Q^{-1}S^{(0)})^\kappa]\ll\sum_{\kappa=1}^{\infty}[(P^{-1}T)^\kappa-(P^{-1}T^{(0)})^\kappa]. \tag{100}$$

But from (100) by virtue of (90) it follows that the development (99) is majorized by (97) and our lemma is proved.

Under the conditions of theorem B in section 4, if the inequalities (56) and (57) are satisfied, we can apply lemma III in replacing the matrix M by Δ_n, $M^{(0)}$ by $\Delta_n^{(0)}$. We obtain

F. *Under the conditions of theorem B, if the inequalities (56) and (57) are satisfied, we have*

$$|A^{-1}-A^{(0)-1}|_p\leqq|\Delta_n^{-1}-\Delta_n^{(0)-1}|_p=\max_{1\leqq\nu\leqq n}(1+M)^{\nu-1}\frac{(1+m)^{n-\nu}-|\Delta_n|}{|\Delta_n|} \qquad (p=1,\infty), \tag{101}$$

where the values and the estimates for the right-hand expression in (101) *are obtained from the formulas* (79) *to* (87°).

Lemma III can be applied to many problems similar to that solved by theorem F. For instance, in the theory of the solutions of linear equations, the following problem has to be dealt with, although its complete discussion is usually avoided:

338

Consider a "triangular" system of linear equations

$$\sum_{\nu=\mu}^{n} a_{\mu\nu}x_\nu = y_\mu \qquad (\mu=1, \ldots, n),$$ (102)

where the coefficients $a_{\mu\nu}$ are only approximate values to the "true" values $b_{\mu\nu}$. Suppose that we have generally

$$|b_{\mu\nu} - a_{\mu\nu}| \leqq \epsilon \qquad (\nu \geqq \mu);$$ (103)

how far is the solution of the system

$$\sum_{\nu=\mu}^{n} b_{\mu\nu}x_\nu = y_\mu$$ (104)

influenced if the $b_{\mu\nu}$ are replaced by $a_{\mu\nu}$?

If we denote the matrix of (102) by $A^{(0)}$ and that of (104) by A, the question can be answered by giving estimates of $|A^{-1} - A^{(0)-1}|_p$ $(p=1, \infty)$. Suppose that we have generally

$$a_{\mu\mu}=1, \quad |a_{\mu\nu}| \leqq M_\mu \qquad (\mu > \nu; \quad \mu=1, \ldots, n-1)$$ (105)

and consider the matrix

$$\Delta(M_1, \ldots, M_{n-1}) = \begin{bmatrix} 1 & -M_1 & \ldots & -M_1 \\ 0 & 1 & \ldots & -M_2 \\ \cdot & \cdot & \ldots & \cdot \\ \cdot & \cdot & \ldots & \cdot \\ \cdot & \cdot & \ldots & \cdot \\ \cdot & \cdot & \ldots & \cdot \\ 0 & 0 & \ldots & 1 \end{bmatrix},$$ (106)

where all elements to the left of the main diagonal are zeros, and all elements to the right of the main diagonal in the μth row are equal to $-M_\mu$.

We obtain from lemma III at once the majorization

$$A^{-1} - A^{(0)-1} \ll \Delta(M_1+\epsilon, \ldots, M_{n-1}+\epsilon)^{-1} - \Delta(M_1, \ldots, M_{n-1})^{-1}.$$

To obtain the inverse of the matrix $\Delta(M_1, \ldots, M_{n-1})$, consider the system of linear equations

$$x_\mu - M_\mu(x_{\mu+1} + \ldots + x_n) = u_\mu \qquad (\mu=1, \ldots, n).$$ (107)

To solve it explicitly, we put

$$s_\mu = \sum_{\kappa=\mu}^{n} x_\kappa \qquad (\mu=1, \ldots, n),$$ (108)

$$1+M_\mu=N_\mu, \qquad N_\mu N_{\mu+1} \ldots N_n = P_\mu \qquad (\mu=1, \ldots, n-1) \qquad N_n=1, \qquad P_n=1.$$ (109)

Then (107) becomes

$$x_\mu = M_\mu s_{\mu+1} + u_\mu \qquad (\mu=1, \ldots, n),$$ (110)

$$s_\mu = N_\mu s_{\mu+1} + u_\mu \qquad (\mu=1, \ldots, n),$$ (111)

where $s_{n+1}=0$. Dividing (111) by P_μ, we obtain

$$\frac{s_\mu}{P_\mu} = \frac{s_{\mu+1}}{P_{\mu+1}} + \frac{u_\mu}{P_\mu}$$

and therefore

$$s_\mu = P_\mu \sum_{\kappa=\mu}^{n} \frac{u_\kappa}{P_\kappa};$$

introducing this in (110) we have

$$x_\mu = u_\mu + M_\mu P_{\mu+1} \sum_{\kappa=\mu+1}^{n} \frac{u_\kappa}{P_\kappa} \qquad (\mu=1,\ldots,n;\ P_{n+1}=0). \tag{112}$$

We obtain therefore for the inverse of our $\Delta(M_1, \ldots, M_{n-1})$, if E denotes the unit matrix,

$$\Delta(M_1, \ldots, M_{n-1})^{-1} = E + DT, \tag{113}$$

where D is the diagonal matrix

$$D = \begin{pmatrix} M_1 P_2 & & & & & \\ & M_2 P_3 & & & & \\ & & \cdot & & \mathbf{0} & \\ & & & \cdot & & \\ & \mathbf{0} & & & \cdot & \\ & & & & & M_{n-1}P_n \\ & & & & & & 0 \end{pmatrix}, \tag{114}$$

and T the triangular matrix

$$T = \begin{pmatrix} 0 & \dfrac{1}{P_2} & \dfrac{1}{P_3} & \cdot & \cdot & \cdot & \dfrac{1}{P_n} \\ & 0 & \dfrac{1}{P_3} & \cdot & \cdot & \cdot & \dfrac{1}{P_n} \\ & & 0 & \cdot & \cdot & \cdot & \dfrac{1}{P_n} \\ & & & \cdot & & & \\ & \mathbf{0} & & & \cdot & & \\ & & & & & \cdot & \\ & & & & & & 0 \end{pmatrix}. \tag{115}$$

The expressions

$$\Delta(M_1+\epsilon, \ldots, M_{n-1}+\epsilon)^{-1} - \Delta(M_1, \ldots, M_{n-1})^{-1} \tag{116}$$

obtained from (113), (114), (115) are rather unwieldy; however, we shall obtain for its norm corresponding to $p=1$ (cf. 66,a) a very simple and elegant expression.

Indeed, if we take the sum of the elements in the νth column of $\Delta(M_1, \ldots, M_{n-1})^{-1}$ and denote it by t'_ν, we have

$$t'_\nu = 1 + \frac{1}{P_\nu} \sum_{\mu=1}^{\nu-1} M_\mu P_{\mu+1} = 1 + \sum_{\mu=1}^{\nu-1} M_\mu \prod_{\kappa=\mu+1}^{\nu-1} (1+M_\kappa).$$

If we now write out $t'_{\nu+1}$, we obtain

$$t'_{\nu+1} = 1 + \left[\sum_{\mu=1}^{\nu-1} M_\mu \prod_{\kappa=\mu+1}^{\nu-1} (1+M_\kappa) \right] (1+M_\nu) + M_\nu;$$

340

and in comparing this with the expression for t'_ν we see that we have $t'_{\nu+1} \equiv (1+M_\nu)t'_\nu$ and therefore

$$t'_\nu = \prod_{\kappa=1}^{\nu-1}(1+M_k). \tag{117}$$

We obtain now for the sum of the elements in the νth column of (116)

$$t_1 = 0, \quad t_2 = \epsilon,$$

$$t_\nu = \prod_{\kappa=1}^{\nu-1}(1+M_\kappa+\epsilon) - \prod_{\kappa=1}^{\nu-1}(1+M_\kappa) \qquad (\nu=1,\ldots,n).$$

In multiplying this by $1+M_\nu+\epsilon$, we obtain

$$\prod_{\kappa=1}^{\nu}(1+M_\kappa+\epsilon) - \prod_{\kappa=1}^{\nu}(1+M_\kappa) - \epsilon\prod_{\kappa=1}^{\nu-1}(1+M_\kappa),$$

and comparing this with $t_{\nu+1}$,

$$t_{\nu+1} = (1+M_\nu+\epsilon)t_\nu + \epsilon\prod_{\kappa=1}^{\nu-1}(1+M_\kappa). \tag{118}$$

Therefore, t_ν is monotonically increasing with ν, and we see that the norm of (116) corresponding to $p=1$ has the value

$$\prod_{\nu=1}^{n-1}(1+M_\nu+\epsilon) - \prod_{\nu=1}^{n-1}(1+M_\nu) \tag{119}$$

and obtain therefore

$$|A^{-1}-A^{(0)-1}|_1 \leqq \prod_{\nu=1}^{n-1}(1+M_\nu+\epsilon) - \prod_{\nu=1}^{n-1}(1+M_\nu) \tag{120}$$

as the solution of our problem. In applications it may be better to use the recurrence formula (118). If all M_μ have the same value M, the expression (113) coincides with that obtained in section 5 for Δ_n^{-1} for $m=0$. But in this case we see from (48) and (49) that the row sums run through the same values as the sums in the columns. We obtain, therefore, in this case the expression

$$(M+1+\epsilon)^{n-1} - (M+1)^{n-1} = (n-1)\epsilon(M+1+\theta\epsilon), \quad 0<\theta<1,$$

as the norm of (116), both for $p=1$ and $p=\infty$.

8. Linear Systems With a Nearly Triangular Matrix

The results of the preceding sections give the means to discuss the following problem concerning the system (1) in the introduction under the conditions (2) and the "triangular" system

$$\sum_{\nu=\mu}^{n}a_{\mu\nu}x_\nu = y_\mu \qquad (\mu=1,\ldots,n), \tag{121}$$

with the matrix $A^{(0)}$. In discussing this problem we can obviously assume that $a_{\mu\mu}=1$ ($\mu=1,\ldots,n$). Then the difference between the solution of (1) and that of (121) is given by the vector $(A^{-1}-A^{(0)-1})\eta$, $\eta=(y_1,\ldots,y_n)$, and the norm of this vector corresponding to one of the indices $p=1, \infty$ does not exceed $|A^{-1}-A^{(0)-1}|_p|\eta|_p$ ($p=1,\infty$), and can indeed for suitable choice of the vector η attain this limit. Therefore, the norms $|A^{-1}-A^{(0)-1}|_p$ measure the error committed in replacing the system (1) by (121).

For an $\epsilon>0, M>0$ being given, how small must $m>0$ be taken in order that we have

$$|A^{-1}-A^{(0)-1}|_p \leqq \epsilon \qquad (p=1,\infty)? \tag{122}$$

341

If we introduce the quantities $|\Delta_n|$ and δ corresponding by (22), (23), and (27) to m and M, we obtain from (79) and (94) the condition

$$(1+M)^{n-1}\frac{\delta}{1-\delta} \leqq \epsilon$$

$$\delta \leqq \frac{\epsilon}{(1+M)^{n-1}+\epsilon} \qquad \left(|\Delta_n| \leqq 1-\frac{m}{M}\right) \tag{123}$$

as long as the condition (75) is satisfied and therefore certainly as long as $M \geqq 1.5/n$ $(n \geqq 4)$.

On the other hand, we have by (30)

$$m = \frac{\delta}{M_n + \theta \dfrac{\delta}{M}}, \tag{124}$$

and from (123) and (124)

$$m \leqq \frac{\epsilon}{M_n(1+M)^{n-1}+\left(M_n+\dfrac{\theta}{M}\right)\epsilon} \qquad \left(|\Delta_n| \leqq 1-\frac{m}{M}\right). \tag{125}$$

We see that it will be sufficient to take

$$m \leqq m_0 \equiv \frac{\epsilon}{M_n(1+M)^{n-1}+\left(M_n+\dfrac{1}{M}\right)\epsilon} \qquad \left(|\Delta_n| \leqq 1-\frac{m}{M}\right) \tag{126}$$

to solve our problem, for instance, if we have $M \geqq 1.5/n$ $(n \geqq 4)$ or $M \geqq 0.5321$ $(n=3)$. For small values of ϵ obviously, only the first term in the denominator is essential, and we have

$$m_0 \approx K(n,M)\epsilon, \quad K(n,M) = \frac{1}{M_n(1+M)^{n-1}}. \tag{127}$$

Tables 1 and 2 give the values of $K(n,M)$ for a set of integer M from 1 to 10 and some values of $M > 1.5/n$. We have obviously

$$K(n,M) < \frac{1}{M^{2n-2}}. \tag{128}$$

TABLE 1. $K(n,M)$

M	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$	$n=8$	$n=9$	$n=10$
1	. 0625	. 0114	. (2)240	. (3)548	. (3)130	. (4)316	. (5)778	. (5)193
2	. 0111	. (2)103	. (3)106	. (4)115	. (5)126	. (6)140	. (7)155	. (8)172
3	. (2)347	. (3)193	. (4)116	. (6)719	. (7)448	. (8)280	. (9)175	. (10)109
4	. (2)143	. (4)526	. (5)206	. (7)821	. (8)328	. (9)131	. (11)524	. (12)210
5	. (3)694	. (4)182	. (6)498	. (7)138	. (9)383	. (10)106	. (12)295	. (14)821
6	. (3)378	. (5)736	. (6)149	. (8)204	. (10)619	. (11)126	. (13)258	. (15)526
7	. (3)223	. (5)336	. (7)522	. (9)815	. (10)127	. (12)199	. (14)211	. (16)486
8	. (3)140	. (5)168	. (7)207	. (9)255	. (11)315	. (13)389	. (15)480	. (17)592
9	. (4)926	. (6)903	. (8)900	. (10)900	. (12)900	. (14)900	. (16)900	. (18)900
10	. (4)636	. (6)515	. (8)424	. (10)351	. (12)290	. (14)239	. (16)198	. (18)164

Table 2. $K(n, M)$

M	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$	$n=8$	$n=9$	$n=10$
.1	------	------	------	------	------	------	------	--------
.2	------	------	------	------	------	.0328	.0197	.0121
.3	------	------	.0866	.0399	.0196	.0101	.(2)533	.(2)289
.4	------	.117	.0438	.0180	.(2)788	.(2)359	.(2)169	.(3)809
.5	.254	.0718	.0241	.(2)891	.(2)349	.(2)142	.(3)592	.(3)252
.6	.181	.0464	.0141	.(2)470	.(2)165	.(3)602	.(3)224	.(4)848
.7	.134	.0313	.00864	.(2)260	.(3)825	.(3)270	.(4)902	.(4)305
.8	.102	.0218	.00548	.(2)150	.(3)431	.(3)127	.(4)382	.(4)116
.9	.0789	.0156	.00359	.(3)894	.(3)233	.(4)623	.(4)169	.(5)462
1.0	.0625	.0114	.00240	.(3)548	.(3)130	.(4)316	.(5)778	.(5)193

M	$n=15$	$n=20$	$n=25$	$n=30$	$n=35$	$n=40$	$n=45$	$n=50$
.1	.0157	.(2)439	.(2)138	.(3)469	.(3)166	.(4)604	.(4)224	.(5)841
.2	.(2)137	.(3)188	.(4)281	.(5)439	.(6)697	.(6)112	.(7)180	.(8)290
.3	.(3)167	.(4)112	.(6)793	.(7)570	.(8)413	.(9)299	.(10)217	.(11)157
.4	.(4)242	.(6)809	.(7)277	.(9)957	.(10)331	.(11)114	.(13)395	.(14)137
.5	.(5)399	.(7)681	.(8)118	.(10)204	.(12)354	.(14)613	.(15)106	.(17)184
.6	.(6)729	.(8)658	.(10)597	.(12)543	.(14)494	.(16)449	.(18)409	.(20)372
.7	.(6)146	.(9)721	.(11)357	.(13)177	.(16)879	.(18)436	.(20)216	.(22)107
.8	.(7)317	.(10)886	.(12)248	.(15)695	.(17)195	.(20)545	.(22)153	.(25)428
.9	.(8)743	.(10)121	.(13)197	.(16)322	.(19)525	.(22)856	.(24)140	.(27)228
1.0	.(8)186	.(11)182	.(14)178	.(17)173	.(20)169	.(23)165	.(26)162	.(29)158

The bound in (126) is obviously the "best" under the condition (75), save that the factor $1/M$ of ϵ in the denominator could be replaced by an (unknown) fraction of it.

If $M<1.5/n$, the use of the general formula (87) is very cumbersome. We can, however, obtain a good working limit for m in the following way. If the inequalities (2) are valid for an $M<1.5/n=M^{(n)}$, then they are also satisfied if M is replaced by $M^{(n)}$, but then the limit for m obtained for $M^{(n)}$ is also sufficient for our M. We obtain therefore in this case the sufficient condition for (122) in the form

$$m \leqq m_0' \equiv \frac{\epsilon}{M_n^{(n)}(1+M^{(n)})^{n-1}+\left(M_n^{(n)}+\dfrac{1}{M_n}\right)\epsilon}. \tag{129}$$

Table 3 gives the values of $K(n, M^{(n)})$ for $n = 1, 2, \ldots, 50$.

TABLE 3. $M = M^{(n)} = 1.5/n$

n	$(1+M)^n - nM - 1$	$K(n, M)$	n	$(1+M)^n - nM - 1$	$K(n, M)$
3	.871	.255	27	1.805	.00755
4	1.074	.134	28	1.811	.00723
5	1.213	.0866	29	1.817	.00694
6	1.315	.0623	30	1.822	.00667
7	1.393	.0480	31	1.827	.00642
8	1.454	.0387	32	1.831	.00619
9	1.504	.0323	33	1.836	.00597
10	1.546	.0276	34	1.840	.00577
11	1.580	.0240	35	1.844	.00558
12	1.610	.0213	36	1.848	.00540
13	1.635	.0190	37	1.851	.00524
14	1.658	.0172	38	1.854	.00508
15	1.677	.0157	39	1.857	.00493
16	1.695	.0144	40	1.860	.00480
17	1.710	.0133	41	1.863	.00466
18	1.724	.0124	42	1.866	.00454
19	1.736	.0116	43	1.869	.00442
20	1.748	.0109	44	1.871	.00431
21	1.758	.0102	45	1.873	.00420
22	1.768	.00965	46	1.876	.00410
23	1.776	.00914	47	1.878	.00401
24	1.784	.00869	48	1.880	.00391
25	1.792	.00827	49	1.882	.00383
26	1.799	.00789	50	1.884	.00374

The expression for m_0' can be written in introducing the value of $M_n^{(n)}$ as

$$\frac{3}{2n} \cdot \frac{\epsilon}{\left(1 + \frac{1.5}{n}\right)^{n-1}\left(\left(1 + \frac{1.5}{n}\right)^n - 2.5\right) + \left(\left(1 + \frac{1.5}{n}\right)^n - 1.5\right)\epsilon}.$$

We observe now that for any positive α and positive x the expression $(1 + x/n)^{n-\alpha}$ monotonically increases and tends to e^x if the positive n increases monotonically to ∞. Indeed, if we put $u = 1/n$, take the logarithm of this expression, differentiate it with respect to u, and multiply by u^2, we obtain

$$\frac{x(u - \alpha u^2)}{1 + ux} - \log(1 + ux)$$

but this expression vanishes for $u = 0$ and decreases for positive u, since its derivative is

$$\frac{-xu}{(1 + ux)^2}(2\alpha + x + \alpha xu).$$

We see, therefore, that

$$\left(1 + \frac{1.5}{n}\right)^{n-1} \uparrow e^{1.5}, \qquad \left(1 + \frac{1.5}{n}\right)^n \uparrow e^{1.5}$$

and our bound for m can therefore be replaced by

$$\frac{3}{2n} \cdot \frac{\epsilon}{e^{1.5}(e^{1.5} - 2.5) + (e^{1.5} - 1.5)\epsilon} > \frac{\epsilon}{(5.93 + 1.99\epsilon)n};$$

we can replace therefore the condition $m \leq m_0'$ by the simpler condition

$$m \leq m_0 = \frac{\epsilon}{n} \frac{1}{6 + 2\epsilon}. \tag{130}$$

On the other hand, we obtain at once the solution of our problem in the case $m = M < 1/(n-1)$ from (87°).

$$m \leqq m_0 \equiv \frac{1}{n-1} \frac{\epsilon}{1+\epsilon} \qquad \left(m = M < \frac{1}{n-1} \right). \tag{131}$$

This is a "best" condition, whereas the condition (130) can still be improved. We see however, in comparing (130) with (131), that the bound in (130) cannot be improved by a factor greater than 6.

From the condition (130) we can finally derive the inequality (7) of the introduction. Indeed, if positive $m_0 < 1/2n$ is given, we can solve (130) with respect to ϵ,

$$\epsilon = \frac{6nm_0}{1-2nm_0},$$

and obtain therefore, in applying (130) for $m = m_0$, the inequality

$$|A^{-1} - A^{(0)\,-1}|_p \leqq \cdot \frac{6nm_0}{1-2nm_0},$$

which holds for all positive $m_0 < 1/2n$ and gives the inequality (7) if we replace here m_0 by m.

9. References

[1] A. Ostrowski, Über die Determinanten mit überwiegender Hauptdiagonale, Comm. Math. Helv. **10,** 69–96 (1937).

[2] A. Ostrowski, Sur les conditions générales pour la régularité des matrices, Rendic. di Mat. e. d. s. applicaz. [V] **10,** 156–168 (1951).

[3] A. Ostrowski, Sur le matrices peu différentes d'une matrice triangulaire, Compt. rend. **233,** 1558–1560 (1951).

[3a] P. Stein and R. L. Rosenberg, On the solution of linear simultaneous equations by iteration, J. London Math. Soc. **23,** 111–117 (1948).

[4] O. Taussky-Todd, A recurring theorem on determinants, Am. Math. Monthly **56,** 672–676 (1949).

WASHINGTON, January 27, 1953.