

Authors: D. J. Evans, M. Hatzopoylos

Title: On the Direct Solution of Preconditioned Linear Equations

Creator: HDML

## ON THE DIRECT SOLUTION OF PRECONDITIONED LINEAR EQUATIONS

M. HATZOPOULOS and D. J. EVANS

A concept of preconditioning hitherto successfully applied to accelerate the convergence rate of iterative methods of solving the large order sparse ill-conditioned linear systems derived from finite difference and finite element analyses of potential problems is now extended to the direct methods of solution of small order dense ill-conditioned matrices. It is shown that by solving the equivalent preconditioned linear system using elimination methods minimises the rounding errors incurred in the computational process. Theoretical analyses justifying this theory are included together with supporting numerical evidence obtained from well known ill-conditioned matrices.

### 1. Introduction

In this paper, we shall be concerned with the solution of ill-conditioned linear systems when the coefficient matrix is of small order and dense.

Several authors [1] have discussed ways of overcoming the problem of ill-conditioning and the main theme of all these methods is to transform the system to one where the coefficient matrix is better conditioned to solve the new system and then transform back to obtain the solution to the original system.

Evans (1967) initially suggested one of these transformations for large sparse matrices, and proved that the coefficient matrix obtained was better conditioned. Then, using the pre-conditioned matrix he was able to develop a class of iterative methods i.e., the Preconditioned Simultaneous Displacement Methods which possessed superior convergence rates.

Similarly, we can extend the preconditioning theory to small order dense ill-conditioned matrices for which the chosen solution method is the direct method of successive (Gaussian) elimination scheme pro-

vided that some form of pivoting is chosen. Following a similar procedure to Evans (1967), we now attempt to transform the given ill-conditioned system to one which is better conditioned and in so doing minimise the rounding errors which are incurred in the solution process.

## 2. Condition numbers and the relative error in the solution

Turing (1948), first recognised that a linear system need possess a «condition number» and that it should depend symmetrically on both  $A$  and  $A^{-1}$  or specifically on the product of their norms. He thus defined, the  $M$  and  $N$  condition numbers as follows:

$$M(A) = n \max_{i,j} |a_{i,j}| \max |a_{i,j}| \quad (1)$$

$$\text{and } N(A) = n^{-1} \left( \sum_{i,j} a_{i,j}^2 \right)^{\frac{1}{2}} \left( \sum_{i,j} a_{i,j} \right)^{\frac{1}{2}}, \quad (2)$$

where terms  $(a_{i,j})$  are the inverse of the matrix  $A = (a_{i,j})$  of order  $n$ .

Von Neumann and Goldstine (1947) defined the  $P$ -condition number of a matrix  $A$  as:

$$P = \frac{|\lambda_1|}{|\lambda_n|} \quad (3)$$

where  $\lambda_1, \lambda_n$  are the eigenvalues of largest and smallest modulus of  $A$  respectively. A significant result which they proved was that if  $A$  is a symmetric positive definite matrix and  $Z$  its computed inverse then, the rounding errors incurred by using the elimination method are bounded and can be expressed as

$$|\mu| \leq 14.24 \cdot P(A) n^2 \varepsilon, \quad (4)$$

where  $\mu$  is the eigenvalue of largest modulus of the error matrix  $I - AZ$ ,  $n$  the order of  $A$  and  $\varepsilon$  the smallest number recognisable to the machine. When  $A$  is non-positive definite the result (4) becomes:

$$|\mu| \leq 36.58 \cdot P(AA^T) n^2 \varepsilon, \quad (5)$$

where:

$$P(AA^T) \geq [P(A)]^2. \quad (6)$$

Similarly the  $k$ -condition number for a general matrix is defined as:

$$K(A) = \sqrt{\frac{\lambda_1(AA^T)}{\lambda_n(AA^T)}} = \sqrt{P(AA^T)}. \quad (7)$$

In addition, the following inequalities can be proved for the condition numbers defined above,

$$\frac{M(A)}{n^2} \leq N(A) \leq M(A) \quad (8)$$

and

$$\frac{M(A)}{n} \leq P(A) \leq nM(A). \quad (9)$$

If we consider small perturbations in the elements to  $A$ , and  $b$  in the given system  $Ax=b$  then we can estimate the error  $\delta x$  on  $x$  from the perturbed system  $(A+\delta A)(x+\delta x) = b+\delta b$ .

It can be proved that the relative error can be expressed as:

$$\frac{\|\delta x\|}{\|x\|} \leq \|A\| \cdot \|A^{-1}\| \cdot \left\{ \frac{\|\delta b\|}{\|A\| \cdot \|\delta A\|} + \frac{\|\delta A\|}{\|A\|} \right\} \cdot \frac{1}{(1 - \|A^{-1}\| \cdot \|\delta A\|)}. \quad (10)$$

If we let  $k = \|A\| \cdot \|A^{-1}\|$ ,  $\delta_1 = \frac{\|\delta b\|}{\|A\| \cdot \|\delta A\|}$ ,  $\delta_2 = \frac{\|\delta A\|}{\|A\|}$  then (10) can be written in the form:

$$\frac{\|\delta x\|}{\|x\|} \leq k(\delta_1 + \delta_2)/(1 - k\delta_1), \quad (11)$$

where it is obvious that  $k$  is an indicator of the relative errors incurred in the computational solution of the linear system.

### 3. The preconditioning method

The system:

$$Ax = b, \quad (12)$$

where  $A$  is a non-singular matrix can without loss of generality be written as:

$$A = I + L + U. \quad (13)$$

where  $I$  is the identity matrix and  $L$  and  $U$  are strictly lower and upper triangular matrices respectively.

Consider now, the non-singular matrices  $(I+\omega L)$  and  $(I+\omega U)$  where  $\omega$  is a parameter to be defined later. Then, the system (12) can be transformed into the equivalent system:

$$(I + \omega L)^{-1} A (I + \omega U)^{-1} [(I + \omega U)x] = (I + \omega L)^{-1} b \quad (14)$$

or

$$B_\omega y = d_\omega, \quad (15)$$

where

$$B_{\omega} = (I + \omega L)^{-1} A (I + \omega U)^{-1} \quad (16)$$

$$d_{\omega} = (I + \omega L)^{-1} b, \quad (17)$$

$$\text{and } y = (I + \omega U)x. \quad (18)$$

By inspection, we see that for  $\omega=0$  the new system (15) reverts back to its original form (12). Thus by introducing the transformation above, we allow  $\omega$  to play the role of a preconditioning parameter such that as  $\omega$  varies in a restricted range  $0 < \omega < W$  (say), a minimum value of the condition number of  $B_{\omega}$  is obtained. Thus, at the optimum  $\omega$  we solve the system (15) and from (18) we calculate the solution  $x$  to the original system.

Clearly the forms of  $(I + \omega L)$  and  $(I + \omega U)$  permit accurate computation of the preconditioned matrix  $B_{\omega}$ , in (16), from simple algorithms of forward and backward substitution on the rows and columns of  $A$  i.e.

$$B_{\omega} = \begin{bmatrix} 1 & & & 0 \\ & \ddots & & \\ & & \ddots & \\ & l_{ij} & & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & & & \\ & \ddots & & a_{ij} \\ & & \ddots & (i < j) \\ & a_{ij} & & 1 \\ & & & \ddots \\ & & & & \ddots \\ & & & & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & \\ & \ddots & & u_{ij} \\ & & \ddots & \\ & 0 & & 1 \end{bmatrix}^{-1}$$

where

$$l_{ij} = \omega a_{ij} \quad i > j, \quad (19)$$

and

$$u_{ij} = \omega a_{ij} \quad i < j. \quad (20)$$

If we let:

$$R = (I + \omega L)^{-1} A \quad (21)$$

then

$$(I + \omega L)R = A,$$

and the algorithm given by:

$$r_{1j} = a_{1j} \quad \text{for } j=1, \dots, n \quad (22)$$

and

$$r_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik} r_{kj} = a_{ij} - \omega \sum_{k=1}^{i-1} a_{ik} r_{kj} \quad i=2, \dots, n$$

whereas the remaining operation to determine  $B_{\omega}$  i.e.,

$$B_{\omega} = R(I + \omega U)^{-1}$$

can either be completed via a similar algorithm to the above working on the transposed system,

$$R^T = (I + \omega U)^T B_\omega^T$$

or by operating on the rows instead of the columns, in which case, we have algorithmic processes given by:

$$R = B_\omega (I + \omega U),$$

or for  $i=1, \dots, n$

$$B_{i1} = r_{i1} \quad (23)$$

and

$$B_{ij} = r_{ij} - \sum_{k=1}^{j-1} B_{ik} u_{kj} = r_{ij} - \omega \sum_{k=1}^{j-1} B_{ik} a_{kj} \quad j=2, \dots, n.$$

Similarly for the right hand side

$$b = (I + \omega L)d$$

the algorithm  $\omega$  is given by:

$$d_1 = b_1 \quad (24)$$

and

$$d_i = b_i - \sum_{k=1}^{i-1} l_{ik} d_k = b_i - \omega \sum_{k=1}^{i-1} a_{ik} d_k \quad i=1, \dots, n.$$

The extra work involved in calculating  $B_\omega$  is  $O(N^3)$  multiplications, however it is worthy of consideration for the improvement in accuracy we obtain for the solution of ill-conditioned systems especially those involving multiple right hand sides. In the above calculations single precision arithmetic with double length accumulation of inner products was used whereas hitherto for ill-conditioned systems double precision working has been employed.

When  $A$  is positive definite  $B_\omega$  can be computed using only  $\frac{h^3}{2}$  operations by writing:

$$B_\omega = S_\omega + S_\omega^T$$

where:

$$S_\omega = (I + \omega L)^{-1} \left( \frac{1}{2} I + L \right) (I + \omega L^T)^{-1}.$$

#### 4. Preconditioning techniques applied to symmetric positive definite matrices

Let  $A$  be a symmetric positive definite matrix and  $D$  the matrix of diagonal entries. Then, we can write  $A$  without loss of generality in the form

$$A = D + L + L^T, \quad (25)$$

and to bring it in line to the form (13), we premultiply and post multiply  $A$  by  $D^{-\frac{1}{2}}$  i.e.,

$$D^{-\frac{1}{2}} A D^{-\frac{1}{2}} = A', \quad (26)$$

which can be written in the form:

$$A' = I + L' + L'^T, \quad (27)$$

where

$$L' = D^{-\frac{1}{2}} L D^{-\frac{1}{2}} \quad \text{and} \quad L'^T = D^{-\frac{1}{2}} L^T D^{-\frac{1}{2}}. \quad (28)$$

This transformation corresponds approximately to a form of scaling or equilibration and improves the condition of the system accordingly. (Forsythe and Moler, 1967).

The pre-conditioned matrix  $B_\omega$  is then given by:

$$B_\omega = (I + \omega L)^{-1} A (I + \omega L^T)^{-1} \quad (29)$$

as derived from (16) where we have written  $L$  for  $L'$  and  $L^T$  for  $L'^T$  for convenience.

Let  $x$  be an eigenvector of the matrix  $B_\omega$  corresponding to the eigenvalue  $\mu$ . Then by definition, we have:

$$(I + \omega L)^{-1} A (I + \omega L^T)^{-1} x = \mu x. \quad (30)$$

If we write:

$$y = (I + \omega L^T)^{-1} x, \quad (31)$$

then (30) can be expressed in the form,

$$Ay = \mu (I + \omega L) (I + \omega L^T) y \quad (32)$$

which on expanding becomes

$$Ay = \mu \left( (1 - \omega) I + \omega A + \omega^2 LL^T \right) y \quad (33)$$

and finally in the more compact form,

$$Ay = \mu \left( \left( 1 - \frac{\omega}{2} \right)^2 I + \omega A + \omega^2 \left( LL^T - \frac{1}{4} I \right) \right) y \quad (34)$$

as given by Andersson (1975). If we now premultiply (34) by the vector  $y^T$ , we have the final result,

$$y^T A y = \mu y^T \left( \left( 1 - \frac{\omega}{2} \right)^2 I + \omega A + \omega^2 \left( LL^T - \frac{1}{4} I \right) \right) y. \quad (35)$$

Since matrix  $A$  is symmetric and positive definite, then it can be shown that there always exists a number  $\alpha$  ( $\neq 0$ ) such that:

$$\frac{y^T y}{y^T A y} \leq \alpha. \quad (36)$$

If we now assume that there exists a number  $b$  for which the relationship:

$$y^T \left( LL^T - \frac{1}{4} I \right) y \leq b y^T A y \quad (37)$$

holds, then from (35), (36) and (37) we have the result:

$$\frac{1}{\mu} \leq \left( 1 - \frac{\omega}{2} \right)^2 \alpha + \omega + \omega^2 b. \quad (38)$$

The previous relationship holds for every eigenvalue of  $B_\omega$  and for the minimum eigenvalue, i.e.  $\mu_{\min}$  we have:

$$\frac{1}{\mu_{\min}} \leq \left( 1 - \frac{\omega}{2} \right)^2 \alpha + \omega + \omega^2 b. \quad (39)$$

Further, it can be shown with a little analysis that the following result is valid, i.e.,

$$(I + \omega L)(I + \omega L^T) = \omega(2 - \omega)A + ((\omega - 1)I + \omega L)((\omega - 1)I + \omega L)^T, \quad (40)$$

which when substituted in (32) yields,

$$A y = \mu(\omega(2 - \omega)A + ((\omega - 1)I + \omega L)((\omega - 1)I + \omega L)^T) y \quad (41)$$

and premultiplying by  $y^T$  we obtain the result,

$$y^T A y = \mu y^T (\omega(2 - \omega)A + ((\omega - 1)I + \omega L)((\omega - 1)I + \omega L)^T) y. \quad (42)$$

However, since by definition, the matrix:

$$((\omega - 1)I + \omega L)((\omega - 1)I + \omega L)^T$$

is positive semi-definite, i.e.

$$y^T ((\omega - 1)I + \omega L)((\omega - 1)I + \omega L)^T y \geq 0, \quad (43)$$

then we have the result

$$y^T A y \geq \mu y^T (\omega(2 - \omega)A) y \quad (44)$$

to be valid for all eigenvalues of the matrix  $B_\omega$  and especially so for the largest one,  $\mu_{\max}$ .

Thus, we have:

$$\mu_{\max} \leq \frac{1}{\omega(2 - \omega)} \quad \omega \in (0, 2). \quad (45)$$

By definition, the  $P$ -condition number of  $B_\omega$  can be expressed as:

$$P(B_\omega) = \frac{\mu_{\max}}{\mu_{\min}} \leq \frac{(2-\omega)\alpha}{4\omega} + \frac{1}{2-\omega} + \frac{\omega b}{2-\omega}. \quad (46)$$

The minimum value of the right hand side of (46) can be shown to be:

$$\sqrt{\frac{1}{2}(2b+1)/\alpha} + \frac{1}{2} \quad (47)$$

and it occurs for an optimum value of  $\omega$  given by:

$$\omega_{\text{op}} = 2 / (1 + \sqrt{2(2b+1)/\alpha}). \quad (48)$$

Thus, we have shown that for the given matrix  $B_\omega$ , as given in (29), then the  $P$  condition number is bounded and satisfies the relationship,

$$P(B_{\omega_{\text{op}}}) \leq \sqrt{\frac{1}{2}(2b+1)\alpha} + \frac{1}{2}, \quad (49)$$

in the  $0 < \omega < 2$ .

**Theorem.** For the symmetric and positive definite matrix  $A = I + L + L^T$  then if:

$$\frac{y^T y}{y^T A y} \leq \alpha \quad \forall y \in R^n, \quad y \neq 0 \quad (50)$$

and

$$y^T \left( LL^T - \frac{1}{4} I \right) y \leq b y^T A y \quad \forall y \in R^n, \quad (51)$$

then the  $P$ -condition number of the pre-conditioned matrix  $B_\omega$  is bounded as in (49) for an optimum value of  $\omega$  given by (48).

For the more general case when  $A = D + L + L^T$  we first use a pre-transformation of the form:

$$D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \quad (52)$$

to recover to given matrix form in (28).

i.e. the preconditioned matrix now has the form,

$$B_\omega = D^{-\frac{1}{2}} (D + \omega L)^{-1} A (D + \omega L^T)^{-1} D^{\frac{1}{2}} \quad (53)$$

and as in the above theorem we can easily prove that:

$$P(B_{\omega_{\text{op}}}) \leq \sqrt{\frac{1}{2}(2b+1)/\alpha} + \frac{1}{2} \quad (54)$$

for:

$$\omega = 2 / (1 + \sqrt{2(2b+1)/\alpha}). \quad (55)$$

If instead of (50) and (51) we use the alternative upper bounds,

$$\frac{y^T D y}{y^T A y} \leq \alpha \quad y \in R^n \quad (56)$$

and

$$y^T \left( L D^{-1} L^T - \frac{1}{4} D \right) y \leq b y^T A y. \quad (57)$$

## 5. Results

**Example 1.** Consider the  $n \times n$  Pascal matrix defined by:

$$A = (a_{ij}) = \begin{cases} \rightarrow k & i=1, j=1, 2, \dots, n \\ \rightarrow k & j=1, i=1, 2, \dots, n \\ \rightarrow a_{i-1, j} + a_{i, j-1} & i, j=2, 3, \dots, n \end{cases} \quad (58)$$

If  $k$  is reciprocal of an integer then the elements of  $A^{-1}$  are integers and:

$$\det(A) = k^n.$$

We consider the  $8 \times 8$  linear system:

$$Ax = b \quad (59)$$

where  $A$  is the Pascal matrix as defined in (58) with  $k=1/7$  and:

$$b(i) = \sum_{j=1}^8 a_{ij}.$$

System (59) is known to have the solution:

$$x^T = (1, 1, 1, 1, 1, 1, 1, 1).$$

The  $P$  and  $N$  condition numbers of (59) are:

$$P(A) = 2.064 \cdot 10^7$$

$$N(A) = 2.583 \cdot 10^6$$

and the relative Euclidean error norm when we solve (59) is:

$$.7720 \cdot 10^{-5}.$$

In the following table we give the  $P$ - and  $N$ -condition numbers and the relative Euclidean error norm versus the pre-conditioning parameter.

TABLE 1

$\omega$	P-Condition Number	N-Condition Number	Relative Euclidean Error Norm
0	$1.524 \cdot 10^6$	$2.679 \cdot 10^5$	$.2367 \cdot 10^{-5}$
.1	$9.501 \cdot 10^5$	$1.376 \cdot 10^5$	$.4736 \cdot 10^{-5}$
.2	$6.214 \cdot 10^5$	$9.519 \cdot 10^4$	$.8558 \cdot 10^{-6}$
.3	$4.264 \cdot 10^5$	$6.885 \cdot 10^4$	$.1669 \cdot 10^{-6}$
.4	$3.074 \cdot 10^5$	$5.137 \cdot 10^4$	$.2010 \cdot 10^{-6}$
.5	$2.321 \cdot 10^5$	$3.925 \cdot 10^4$	$.1815 \cdot 10^{-6}$
.6	$1.817 \cdot 10^5$	$3.056 \cdot 10^4$	$.1803 \cdot 10^{-6}$
.7	$1.460 \cdot 10^5$	$2.414 \cdot 10^4$	$.9982 \cdot 10^{-6}$
.8	$1.195 \cdot 10^5$	$1.932 \cdot 10^4$	$.9950 \cdot 10^{-7}$
.9	$9.932 \cdot 10^4$	$1.563 \cdot 10^4$	$.7923 \cdot 10^{-7}$
1	$8.356 \cdot 10^4$	$1.278 \cdot 10^4$	$.9668 \cdot 10^{-7}$
1.1	$7.118 \cdot 10^4$	$1.057 \cdot 10^4$	$.9328 \cdot 10^{-7}$
1.2	$6.151 \cdot 10^4$	$8.872 \cdot 10^3$	$.7639 \cdot 10^{-7}$
1.3	$5.418 \cdot 10^4$	$7.597 \cdot 10^3$	$.9334 \cdot 10^{-7}$
1.4	$4.908 \cdot 10^4$	$6.702 \cdot 10^3$	$.9158 \cdot 10^{-7}$
1.5	$4.644 \cdot 10^4$	$6.190 \cdot 10^3$	$.8686 \cdot 10^{-7}$
1.6	$4.702 \cdot 10^4$	$6.137 \cdot 10^3$	$.9859 \cdot 10^{-7}$
1.7	$5.280 \cdot 10^4$	$6.776 \cdot 10^3$	$.1319 \cdot 10^{-6}$
1.8	$6.877 \cdot 10^4$	$8.719 \cdot 10^3$	$.1696 \cdot 10^{-6}$
1.9	$1.081 \cdot 10^5$	$1.361 \cdot 10^4$	$.2997 \cdot 10^{-5}$
2	$2.074 \cdot 10^5$	$2.663 \cdot 10^4$	$.1515 \cdot 10^{-5}$

Comparing the condition numbers with those of the original system we see that they decrease by a factor of  $10^5$  and that the error decreased by a factor of  $10^8$ .

Let us now for  $n = 4, 5, \dots, 12$  consider the Pascal matrices with  $k = 1/7$ . In the following table we give their  $P$ -condition number without preconditioning and for optimum pre-conditioning.

TABLE 2

Order of Matrix	P-Condition Number without Pre-Conditioning	P-Condition Number for optimum Pre-Conditioning
4	$6.919 \cdot 10^2$	2.823.10
5	$8.517 \cdot 10^3$	$1.548 \cdot 10^2$
6	$1.107 \cdot 10^5$	$9.724 \cdot 10^2$
7	$1.777 \cdot 10^6$	$6.523 \cdot 10^3$
8	$2.064 \cdot 10^7$	$4.644 \cdot 10^4$
9	$2.907 \cdot 10^8$	$3.408 \cdot 10^5$
10	$4.154 \cdot 10^9$	$2.548 \cdot 10^6$
11	$6.001 \cdot 10^{10}$	$1.952 \cdot 10^7$
12	$8.786 \cdot 10^{11}$	$1.527 \cdot 10^8$

We see that the P-condition number decreases by a factor of  $10^3$ .

**Example 2.** In Table 3 we present the P and N condition numbers of the Wilson Matrix, i.e.,

$$\begin{pmatrix} 5 & 7 & 6 & 5 \\ 7 & 10 & 8 & 7 \\ 6 & 8 & 10 & 9 \\ 5 & 7 & 9 & 10 \end{pmatrix}$$

which possesses the P-condition number =  $2.9840 \cdot 10^3$ .

TABLE 3

$\omega$	P-Condition	N-Condition
0	$2.2244 \cdot 10^3$	565.72
.2	$1.1651 \cdot 10^3$	298.33
.4	$6.7699 \cdot 10^2$	177.73
.6	$4.6067 \cdot 10^2$	123.43
.8	$3.7357 \cdot 10^2$	100.85
.9	$3.5858 \cdot 10^2$	<b>96.95</b>
1	<b><math>3.5855 \cdot 10^2</math></b>	97.08
1.1	$3.7303 \cdot 10^2$	101.18
1.2	$4.0337 \cdot 10^2$	109.95
1.4	$5.2866 \cdot 10^2$	143.14
1.6	$8.0630 \cdot 10^2$	213.84
1.8	$1.4049 \cdot 10^3$	362.01
2	$2.6869 \cdot 10^3$	679.20

For comparison, we give the error matrix  $|I-AA^{-1}|$  without preconditioning and for  $\omega=1$

without preconditioning

$$\begin{bmatrix} .256.10^{-8} & .175.10^{-8} & .145.10^{-8} & .553.10^{-9} \\ .142.10^{-8} & .768.10^{-9} & .104.10^{-8} & .324.10^{-9} \\ .246.10^{-8} & .179.10^{-8} & .147.10^{-8} & .521.10^{-9} \\ .219.10^{-8} & .165.10^{-8} & .137.10^{-8} & .507.10^{-9} \end{bmatrix}$$

$\omega=1$

$$\begin{bmatrix} .867.10^{-18} & .867.10^{-18} & .325.10^{-18} & .542.10^{-18} \\ .132.10^{-10} & .164.10^{-10} & .704.10^{-11} & .424.10^{-11} \\ .168.10^{-9} & .116.10^{-9} & .865.10^{-10} & .288.10^{-10} \\ .289.10^{-9} & .220.10^{-9} & .109.10^{-9} & .305.10^{-10} \end{bmatrix}$$

**Example 3.** Finally we give an example of a non-symmetric matrix. Here instead of the  $P$ -condition number we have to minimise the  $k$ -condition number. We consider non-symmetric  $6 \times 6$  Vanthermonde matrix:

$$V(x_1, x_2, x_3, x_4, x_5, x_6) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \\ x_1^2 & x_2^2 & x_3^2 & x_4^2 & x_5^2 & x_6^2 \\ x_1^3 & x_2^3 & x_3^3 & x_4^3 & x_5^3 & x_6^3 \\ x_1^4 & x_2^4 & x_3^4 & x_4^4 & x_5^4 & x_6^4 \\ x_1^5 & x_2^5 & x_3^5 & x_4^5 & x_5^5 & x_6^5 \end{bmatrix}$$

and we set:

$$x_1 = 1 \quad x_2 = 2 \quad x_3 = 3 \quad x_4 = 4 \quad x_5 = 5 \quad x_6 = 6.$$

The condition numbers of  $A$  are:

$$K(A) = 7.311.10^5$$

$$N(A) = 1.220.10^6$$

$$M(A) = 1.975.10^6$$

In the following table we give the values of these condition numbers versus  $\omega$ .

TABLE 4

$\omega$	K-condition Number	N-condition Number	M-condition Number
0	$7.581 \cdot 10^3$	$1.399 \cdot 10^3$	$7.876 \cdot 10^3$
.1	$5.247 \cdot 10^3$	$1.017 \cdot 10^3$	$4.790 \cdot 10^3$
.2	$3.702 \cdot 10^3$	$7.588 \cdot 10^2$	$3.953 \cdot 10^3$
.3	$2.666 \cdot 10^3$	$5.781 \cdot 10^2$	$3.583 \cdot 10^3$
.4	$1.956 \cdot 10^3$	$4.478 \cdot 10^2$	$3.230 \cdot 10^3$
.5	$1.477 \cdot 10^3$	$3.514 \cdot 10^2$	$2.896 \cdot 10^3$
.6	$1.156 \cdot 10^3$	$2.786 \cdot 10^2$	$2.580 \cdot 10^3$
.7	$9.338 \cdot 10^2$	$2.228 \cdot 10^2$	$2.283 \cdot 10^3$
.8	$7.703 \cdot 10^2$	$1.796 \cdot 10^2$	$2.003 \cdot 10^3$
.9	$6.455 \cdot 10^2$	$1.458 \cdot 10^2$	$1.742 \cdot 10^3$
1	$5.481 \cdot 10^2$	$1.194 \cdot 10^2$	$1.499 \cdot 10^3$
1.1	$4.715 \cdot 10^2$	$9.885 \cdot 10$	$1.290 \cdot 10^3$
1.2	$4.121 \cdot 10^2$	$8.305 \cdot 10$	$1.135 \cdot 10^3$
1.3	$3.689 \cdot 10^2$	$7.152 \cdot 10$	$9.805 \cdot 10^2$
1.4	$3.441 \cdot 10^2$	$6.432 \cdot 10$	$8.818 \cdot 10^2$
1.5	$3.456 \cdot 10^2$	$6.245 \cdot 10$	$8.267 \cdot 10^2$
1.6	$3.914 \cdot 10^2$	$6.866 \cdot 10$	$7.715 \cdot 10^2$
1.7	$5.181 \cdot 10^2$	$8.889 \cdot 10$	$7.164 \cdot 10^2$
1.8	$7.955 \cdot 10^2$	$1.347 \cdot 10^3$	$7.694 \cdot 10^2$
1.9	$1.553 \cdot 10^3$	$2.277 \cdot 10^3$	$1.565 \cdot 10^3$
2	$2.428 \cdot 10^3$	$4.677 \cdot 10^3$	$3.018 \cdot 10^3$

## 6. Pre-conditioning applications whilst maintaining integer coefficients

It can be argued that pre-conditioning as it has been developed in the last section does not maintain integer coefficients of a linear system, if they happen to be so. We now introduce a similar pre-conditioning that maintains integer coefficients.



- [3] G. E. Forsythe and Cleve P. Moler: *Computer Solution of Linear Algebraic Systems*. Prentice Hall Inc., Englewood Cliffs, N. J. (1967).
- [4] D. J. Evans: *The Use of Pre-conditioning in Iterative Methods for Solving Linear Equations with Symmetric Positive Definite Matrices*. *J. Inst Maths.* (1967) **4**, 295—314.
- [5] D. J. Evans: *Iterative Sparse Matrix Algorithms*. Software for Numerical Mathematics, ed. D. J. Evans, Loughborough University Conference Proceedings, Acad. Press (1974).
- [6] D. J. Evans: *Comparison of the Convergence Rates of Iterative Methods for Solving Linear Equations*. Presented at the Caratheodory Centenary Symposium, Greek Mathematical Society, Athens (1973).
- [7] D. J. Evans, M. Hatzopoulos: *A Note on the LU factorisation of certain matrices*. *C.A.C.M.* p. 278, Vol. **18** (1975).
- [8] M. Hatzopoulos: *The Use of Pre-Conditioning for Solving Ill-Conditioned Linear Equations*: M. Sc. Thesis, Loughborough University (1972).
- [9] M. Hatzopoulos: *Pre-conditioning and other computational techniques for the direct solution of linear systems*. Ph. D. Thesis, Loughborough University (1974).
- [10] Von Neumann and H. H. Goldstine: *Numerical Inverting of Matrices of High Order*, *Bull. Amer. Math. Soc.* **53** (1947), 1021—1099.
- [11] J. Todd: *On Condition Numbers*. Colloques Internationaux de Centre National de la Recherche Scientifique, B. Besancon 7—14 Septembre (1966).
- [12] J. Wilkinson: *Rounding off Errors in Algebraic Processes*. Notes on Applied Science No 32 N.P.L.
- [13] J. Wilkinson: *The solution of Ill-Conditioned Linear Systems*. *Mathematical Methods for Digital Computers Vol. 2* Edited by A. Ralston and M. Wilf. John Wiley and Sons, Inc. (1967).

Department of Computer Studies  
University of Technology  
Loughborough, Leicestershire  
England

(Received September 7, 1975)