MAX-PLANCK-INSTITUT FÜR PLASMAPHYSIK

GARCHING BEI MÜNCHEN

A CONSTRUCTIVE METHOD OF SOLVING THE

LIAPOUNOV EQUATION FOR COMPLEX

HESSENBERG MATRICES

R. Meyer-Spasche

IPP 6/106   R. Meyer-Spasche   A Constructive Method
of Solving the Liapounov
Equation for Complex
Hessenberg-Matrices

(in English)

Abstract

This paper describes a method of solving the Liapounov
equation (1) $HM + M^*H = 2D$, M in upper Hessenberg form,
D diagonal. Initialising the first row of the matrix A
arbitarily, one can find (by solving equations with one
unknown) the unknown elements of A such that
(2) $AM + M^*A^* = 2F$, where A differs from a Hermitian matrix
only in that its diagonal elements need not be real. F is
a diagonal matrix which is uniquely determined by the first
row of A. By solving equation (2) for several initial values
one may generate several matrices A and F (in the most un-
favourable case 2n-1 A's and F's are needed) and superpose
them to get n linearly independent Hermitian matrices H_j
and D_j respectively for which $H_j M + M^* H_j = 2 D_j$ is valid.
Then one can solve the real system $\sum_{j=1}^{n} P_j D_j = D$ to obtain
the solution $H := \sum_{j=1}^{n} P_j \cdot H_j$ of eq.(1).

1. Introduction

In [1] it is shown how the stability problem can be solved for real n x n matrices. How this method can be extended to complex matrices, is described in [3]. The program is running with satisfactory results, even for matrices with orders ⩾ 50.

A central point of this method is the solution of the Liapounov equation

$$HM + M^*H = 2D,$$

where M is a matrix in special upper Hessenberg form, which is similar to the matrix $\widetilde{M}$, given initially, and D is diagonal.

When a Hermitian solution H has been found and D is positive definite, H has the same inertia as $\widetilde{M}$. (see [2], Theorem 1).

2. Preliminaries

Let $\mathbb{R}$ be the field of real numbers, $\mathbb{C}$ the field of complex numbers. Let $M \in L(\mathbb{C}^n)$, i.e. a complex n x n matrix, and let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of M. The following theorem is then valid ([2], § 4, p. 78):

Theorem 1: If $\Delta(M) := \prod\limits_{i,j=1}^{n} (\lambda_i + \overline{\lambda}_j) \neq 0$ and P is a given Hermitian positive-definite matrix, then there exists an unique H satisfying $HM + M^*H = P$, and H is Hermitian.

Remark: $HM + M^*H = 0$ has only the trivial solution $H = 0$ iff $\Delta(M) \neq 0$ is valid ([2], § 2, p. 75).

For the sake of simplicity we will specially choose $P = 2D$, D diagonal, in the following. D is then real because $D^* = D$. The following definitions and lemmas are direct generalizations of the results for real matrices in [1], § 2.

Definition: A Hermitian matrix H and a real diagonal matrix D are called <u>a Liapounov pair with respect to M</u> if they satisfy the equation

(1) $$HM + M^*H = 2D.$$

Lemma 1: Two matrices H and D, $H = H^*$, $D = D^*$ diagonal, form a Liapounov pair with respect to M iff there is a complex matrix T, $T^* = -T$, such that $HM = T + D$.

Proof:

1) $HM = T + D \Rightarrow HM + M^*H = T + D + (T + D)^* = 2D$ because $T + T^* = 0$. H,D form thus a Liapounov pair.

2) Let $HM + M^*H = 2D$, $T := HM - D \Rightarrow T^* = M^*H - D^*$, and $T + T^* = HM - D + M^*H - D = 0$.

Lemma 2: Let $\triangle (M) \neq 0$. If $\{H_j, D_j\}$, $j = 1, \ldots, m$,

are Liapounov pairs with respect to M, then

if $H_1, \ldots H_m$ are linearly independent in $\mathbb{R}$, so too are

$D_1, \ldots D_m$.

Proof: Let us assume that $D_1, \ldots D_m$ are linearly

dependent, i.e. there are $a_1, \ldots, a_m \in \mathbb{R}$, $(a_1, \ldots, a_m)$

$\neq (0, \ldots, 0)$, such that $\sum_{j=1}^{m} a_j D_j = 0$. One then has

$2 \sum_{j=1}^{m} a_j D_j = \sum_{j=1}^{m} a_j (H_j M + M^* H_j) = (\sum_{j=1}^{m} a_j H_j) M + M^* (\sum_{j=1}^{m} a_j H_j)$

$= 0$. According to the remark following theorem 1 this

means, however, that the H_j are linearly dependent.

Lemma 2 has thus been proved.

We will now confine our attention to marices M in special

upper Hessenberg form (i.e. $m_{ij} = 0$ if $i > j + 1$ and $m_{i+1, i}$

$\neq 0$; $i, j = 1, \ldots, n$). For the solution of the stability

problem, this is no constraint (see [1], p.1 f or [3], § 3).

If one wants to solve the Liapunov equation for other purposes,

the situation is quite different:

Let S be a unitary matrix such that $S \tilde{M} S^* = M$ is in upper

Hessenberg form (for instance S may be constructed by

Householder's method).

If all elements in the subdiagonal of M are non-zero, then

the solution can simply be transformed back: Let $HM + M^* H = 2D$,

then $\tilde{H}: = S^* H S$ satisfies

$$\tilde{H} \tilde{M} + \tilde{M}^* \tilde{H} = S^* (HM + M^* H) S = 2 S^* D S.$$

But if some element in the subdiagonal of M is zero, M has

to be split into blocks.

$$\begin{pmatrix} M_1 & \vdots & M_2 \\ \hdashline 0 & \vdots & M_3 \end{pmatrix}$$

M_1 and M_3 are upper Hessenberg matrices, for which the Liapunov equation can ce solved. But if $M_2 \neq 0$, these two solutions do not induce the solution of the complete problem in a simple manner. In this case the method is not applicable.

In the following we shall reduce the solution of the Liapounov equation (1) to repeated solution of the equation

(2) $HM = T + D$, $T^* = -T$, $D^* = D$ diagonal.

For this purpose the solvability of eq. (2) has to be studied more closely.

<u>Lemma 3:</u> Let M be an upper Hessenberg matrix with non-zero elements in the lower co-diagonal, and let
$h := (a_1, a_2 + ib_2, \ldots, a_n + ib_n)$. There then exists an unique complex matrix A with the following properties:

(3.1) A contains h as the first row

(3.2) $A = B + iC$, $B^* = B$, $C = \text{diag}(0, c_2, \ldots c_n) \in L(\mathbb{R}^n)$.

(3.3) $AM = T + D$, $T^* = -T$, $D^* = D$ diagonal matrix.

Proof: Let $A = (a_{jk})$, $B = (b_{jk})$, $AM = (am_{jk})$. As the first

row of A is given, we can calculate the first row of

$AM = T + D$. Since B is Hermitian and T anti-Hermitian, the

first columns of A and AM are then known, hence a_{21} and am_{21}

in particular. a_{22} is now calculated from

$$a_{21}m_{11} + a_{22}m_{21} = am_{21} = - \overline{am_{12}}.$$

The other elements in the second column of A are calculated

accordingly. By means of (3.2) it is now possible to use

the second column of A to complete the second row of A,

etc.

If the $(j - 1)$-th row of A is known, $j = 2, \ldots n$, the

j-th column of A is generally obtained from

$$(3) \quad a_{\ell j} = - \frac{1}{m_{j,j-1}} \left(\sum_{k=1}^{j-1} a_{\ell k} m_{k,j-1} + \overline{am_{j-1,\ell}} \right), \quad \ell = j, \ldots, n$$

since all the elements on the r.h.s. of eq.(3) are known

and it was postulated that $m_{j,j-1} \neq 0$, $j=2,\ldots,n$. The j-th

row of A can then be determined.

At no stage of the calculations are known quantities

changed. One thus obtains uniquely determined A, T, and D,

which by virtue of construction have the required properties.

This then allows the basic theorem of the method to be

proved:

Theorem 2:   If $M \in L(\mathbb{C}^n)$ is a matrix in upper Hessenberg

form with non-zero elements in its lower co-diagonal and

if $\Delta(M) \neq 0$, there are at least n in $\mathbb{R}$ linearly independent

Liapounov pairs with respect to M.

Proof: Let $h_1, \ldots, h_n$ be linearly independent vectors from

$\mathbb{C}^n$ with vanishing imaginary parts and $h_{n+1}, \ldots, h_{2n-1}$

linearly independent vectors from $\mathbb{C}^n$ whose real parts and

first components vanish.

For convenience the vectors (1, 0, ..., 0), ...,

(0, ..., 0, 1), (0, i, 0....0),...,(0, ..., 0, i) will be

selected.

By the method just developed one obtains matrices $A_1, \ldots A_{2n-1}$,

where A_j contains h_j as its first row and has the properties

(3.2) and (3.3). The object now is to superpose these A_j in

such a way that n linearly independent matrices with real

diagonal elements are obtained.

One gets $A_j = B_j + iC_j$, $C_j = \text{diag} (0, c_{j2}, \ldots, c_{jn}) \in L(\mathbb{R}^n)$,

$c_j := (c_{j2}, \ldots, c_{jn})$, $j = 1, \ldots, 2n-1$.

$c_j, c_{n+1}, \ldots, c_{2n-1}, j \in \{1, \ldots, n\}$, are n (n - 1)-dimen-

sional real vectors and are thus linearly dependent. There

are thus real numbers $a_1^{(j)}, \ldots, a_n^{(j)} \in \mathbb{R}$ such that

$(a_1^{(j)}, \ldots, a_n^{(j)}) \neq (0, \ldots, 0)$ and

(4) $a_1^{(j)} c_j + \sum_{k=1}^{n-1} a_{k+1}^{(j)} c_{n+k} = 0$

If for some j $a_1^{(j)}$ is equal to zero in every such representation, there is some k_o, $2 \leq k_o \leq n - 1$, with $a_{k_o}^{(j)} \neq 0$. In this case one may permute the subscripts of $a_1^{(j)} c_j$ and $a_{k_o+1}^{(j)} c_{n+k_o}$. Accordingly, the subscripts of the related A,T and D must be permuted, these permutations being maintained in all subsequent calculations. $a_1^{(j)}$ can thus always be set equal to unity.

$$H_j := A_j + \sum_{k=1}^{n-1} a_{k+1}^{(j)} A_{n+k}, \quad j = 1, \ldots, n$$ are then Hermitian matrices which

(a) are linearly independent since their first rows are linearly independent and

(b) solve eq.(2) since

$$H_j M = A_j M + \sum_{k=1}^{n-1} a_{k+1}^{(j)} A_{n+k} M$$

$$= T_j + D_j + \sum_{k=1}^{n-1} a_{k+1}^{(j)} (T_{n+k} + D_{n+k})$$

$$= \underbrace{T_j + \sum_{k=1}^{n-1} a_{k+1}^{(j)} T_{n+k}}_{=:\ T^{(j)}} + \underbrace{D_j + \sum_{k=1}^{n-1} a_{k+1}^{(j)} D_{n+k}}_{=:\ D^{(j)}}$$

$T^{(j)\ast} = -T^{(j)}$, $D^{(j)\ast} = D^{(j)}$, since the $a_k^{(j)}$ are all real.

In accordance with lemma 2 the $D^{(j)}$, $j = 1, \ldots, n$, are then also linearly independent. Theorem 2 has thus been proved.

3. The method

The Liapounov equation (1) is now solved in the following way:

By the method developed in the proof of theorem 2 one can generate n Liapounov pairs $\{H_j, D^{(j)}\}$, $j=1, \ldots, n$.

If $\Delta(M) \neq 0$, the linear system

$$(5) \qquad \sum_{j=1}^{n} p_j D^{(j)} = D$$

is nonsingular and thus has an unique solution $(p_1, \ldots, p_n)$. $H := \sum_{j=1}^{n} p_j H_j$ is then the required solution of eq.(1) because

$$HM + M^* H = \sum_{j=1}^{n} p_j H_j M + M^* \sum_{j=1}^{n} p_j H_j$$

$$= \sum_{j=1}^{n} p_j (T^{(j)} + D^{(j)}) + \sum_{j=1}^{n} p_j (-T^{(j)} + D^{(j)})$$

$$= 2 \sum_{j=1}^{n} p_j D^{(j)}$$

$$= 2 D.$$

If $\Delta(M) = 0$, two cases may occur:

1) $D^{(1)}, \ldots, D^{(n)}$ are linearly dependent, but (5) has some solution $(p_1, \ldots, p_n)$ Then H: $= \sum\limits_{j=1}^{n} p_j D^{(j)}$ is a (not uniquely determined) solution of eq. (1).

2) (5) has no solution. Then eq. (1) has no solution too.

As the following example shows, this method is more cumbersome for small n than the known direct methods of solving eq.(1). If, however, n is large (e.g. n = 50), it is more advantageous to solve only systems of equations of order $\leqslant$ n than a system of equations of order $n^2 - n$. Number of equations to be solved:

(1) In the most favourable case (the diagonals of the first n calculated A_j are real):

Equation (2) has to be solved n times and a system of equations of order n once.

(2) In the most unfavourable case (the diagonals of all A_j contain non-real elements):

Equation (2) has to be solved (2n-1) times and (n - 1) systems of equations of order n - 1, and a system of equations of order n have to be solved.

Solving eq.(2) once means solving $(n-1)+(n-2)+1 = \frac{n(n-1)}{2}$

equations with one unknown.

Examples:

Let us illustrate the method by two examples, a favourable

one and a more unfavourable one. The first one shows that

permutations of subscripts may be necessary.

1) Let $M := \begin{pmatrix} 1 & i \\ i & -2 \end{pmatrix}$, $D := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

The eigenvalues of M are given by $\lambda_{1,2} = -\frac{1}{2} \pm \sqrt{\frac{5}{4}}$

hence $\Delta(M) \neq 0$

$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $A_2 = \begin{pmatrix} 0 & 1 \\ 1 & -i \end{pmatrix}$, $A_3 = \begin{pmatrix} 0 & i \\ -i & -1 \end{pmatrix}$.

$H_1 = A_1$, $H_2 = A_3$, $D_1 = \text{diag}(1, -2)$, $D_2 = \text{diag}(-1, 3)$.

$P_1 D_1 + P_2 D_2 = I \Rightarrow P_1 = 4$, $P_2 = 3$

$H = P_1 H_1 + P_2 H_2 = \begin{pmatrix} 4 & 3i \\ -3i & 1 \end{pmatrix}$.

2) Let $M = \begin{pmatrix} 1 & i \\ 1 & 1 \end{pmatrix}$, $D = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\lambda_{1,2} = 1 \pm e^{i\pi/4}$

The A_j are computed to

$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$, $A_2 = \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix}$, $A_3 = \begin{pmatrix} 0 & i \\ -i & 2i \end{pmatrix}$,

$D_1 = \text{diag}(1, 0)$, $D_2 = \text{diag}(1, -2)$, $D_3 = \text{diag}(0, 1)$

$a_1^{(1)} = 1$, $a_2^{(1)} = -\frac{1}{2}$; $a_1^{(2)} = 1$, $a_2^{(2)} = 0$

$H_1 = a_1^{(1)} A_1 + a_2^{(1)} A_3 = \begin{pmatrix} 1 & -\frac{1}{2}i \\ \frac{1}{2}i & 0 \end{pmatrix}$, $D^{(1)} = a_1^{(1)} D_1 + a_2^{(1)} D_3 = \begin{pmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}$

$$H_2 = a_1^{(2)} A_2 = \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix} \qquad D^{(2)} = D_2$$

$$P_1 D^{(1)} + P_2 D^{(2)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow P_1 = 2, \ P_2 = -1.$$

$$H = P_1 H_1 + P_2 H_2 = \begin{pmatrix} 2 & -1-i \\ -1+i & 2 \end{pmatrix}.$$

References

[1] Howland, J.L., Senez, J.A.: A Constructive Method
for the Solution of the Stability Problem.
Numer. Math. 16, 1-7 (1970)

[2] Ostrowski, A., Schneider, H.: Some Theorems on
the Inertia of General Matrices.
J. Math. Anal. Appl. 4, 72-84 (1962)

[3] Meyer-Spasche, R: Ein Verfahren zur Lösung des
Stabilitätsproblems für komplexe Matrizen (A Method
for Solving the Stability Problem for Complexe
Matrices) IPP 6/104, 1972.