Linear Difference Equations and Matrix Theorems

Germund Dahquist
Randall LeVeque

Dahlquist's Lecture Notes from the course
Numerical Treatment of Initial Value Problems
at the Royal Institute of Technology (KTH), Stockholm

(Translated by RJL while a visiting student at KTH in 1981)

Germund Dahlquist and Randall LeVeque

LINEAR DIFFERENCE EQUATIONS AND MATRIX THEOREMS

Lecture Notes in the course of
Numerical Treatment of Initial Value Problems
at the Royal Institute of Technology, Stockholm

The intention is that these note will successively be supplemented, e.g. by exercises.

March 1981

CONTENTS

1. Systems with constant coefficients

In the study of numerical methods for differential equations, we need a simple test problem for which we can easily compare the solution which the numerical method gives with the exact solution of the problem. This often leads to considering a linear system of difference equations with constant coefficients. Since a difference equation of higher order can always be reduced to a system of first order, we first consider systems of first order, which we write in matrix-vector form as

$$y_{n+1} = Ay_n + g_n \qquad (n = 0,1,2,\ldots), \tag{1.1}$$

where $\{g_n\}$ is a given sequence of vectors in $\mathbb{C}^k$, A is a constant $k\times k$ matrix, $y_0 \in \mathbb{C}^k$ is a given initial vector and $y_1, y_2, \ldots$ are recursively defined by the difference equation.

By induction we find that the solution is

$$y_n = A^n y_0 + \sum_{j=1}^{n} A^{n-j} g_{j-1}. \tag{1.2}$$

It is of particular interest to determine the behavior of y_n as $n\to\infty$ in the homogeneous case (i.e. $g_j = 0\ \forall j$). If A is diagonalizable, define z by the coordinate transformation $y = Vz$, where V is a $k\times k$ matrix whose columns are the eigenvectors of A. The components of z are thus the coefficients in the representation of y as a linear combination of the eigenvectors of Denote by λ_i the eigenvalues of A. Then

$$V^{-1}AV = D := \operatorname{diag}(\lambda_i). \tag{1.3}$$

From (1.1) it follows that *in the homogeneous, diagonalizable case*

$$z_{n+1} = V^{-1}AVz_n = Dz_n.$$

The vector equation therefore reduces to k scalar equations

$$z_{n+1}^{(i)} = \lambda_i z_n^{(i)} \qquad (i = 1,2,\ldots,k),$$

where $z^{(i)}$ is the i^{th} component of the vector z. The solution is simply

$$z_n^{(i)} = \lambda_i^n z_0^{(i)},$$

and hence

$$y_n = VD^nV^{-1}y_0.$$

We find therefore that *in the diagonalizable case*,

(i) $y_n \to 0$ as $h \to \infty$ for every initial vector y_0 iff $|\lambda_i| < 1$ for $i = 1,2,\ldots,k$.

(ii) y_n is bounded as $n \to \infty$ for every initial vector y_0 iff $|\lambda_i| \leq 1$ for $i = 1,2,\ldots,k$.

In the general case, *when A is not diagonalizable, the result* (ii) *must be modified somewhat*. Since $y_n = A^n y_0$, statements about the behavior of y_n for arbitrary initial vectors y_0 are equivalent to statements about the behavior of A^n. A matrix is diagonalizable if and only if it has no *defective* eigenvalues.

Definition. An eigenvalue is called defective if the number of linearly independent eigenvectors corresponding to it is less than the eigenvalue's multiplicity.

Since every eigenvalue has at least one eigenvector, only multiple eigenvalues can be defective. Not all multiple eigenvalues are defective, however.

Example. Consider

$$A = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}.$$

If $a = 0$, every vector is an eigenvector and so the double eigenvalue 1 is not defective. If $a \neq 0$, then there is only one eigenvector, namely $[1,0]^T$. In this case the double eigenvalue 1 is defective.

■

When A has defective eigenvalues we can no longer diagonalize A as in (1.3). There is, however, a standard generalization of this decomposition called the Jordan Canonical Form. To this end we first introduce the *shift matrix* of order g,

$$S_q = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}, \quad (q > 1), \quad S_1 = 0.$$

so called because (for $q > 1$),

$$S_q[x_1, x_2, \ldots, x_q]^T = [x_2, x_3, \ldots, x_q, 0]^T.$$

It is easy to see that $S_q^q = 0$ and that S_q has 0 as a q-fold eigenvalue with $[1,0,\ldots,0]^T$ as its only eigenvector.

If I_q denotes the identity matrix of order q, then a matrix of the form $\lambda I_q + S_q$ is called a *Jordan block* of order q. For example, Jordan blocks of orders 1, 2 and 3 are

$$\lambda, \quad \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} \text{ and } \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}.$$

A Jordan block has λ as a q-fold eigenvalue with the single eigenvector given above. The eigenvalue in a Jordan block is thus defective if $q > 1$.

Lemma 1.1. Jordan Canonical Form (JCF):
Every k×k matrix A can be transformed to the form $T^{-1}AT = J$, where

$$J = \begin{bmatrix} J_1 & & & O \\ & J_2 & & \\ & & \ddots & \\ O & & & J_p \end{bmatrix}$$

with each $J_i = \lambda_i I_{q_i} + S_{q_i}$ a Jordan block of order q_i. The same eigenvalue c occur in several blocks. A and J have the same eigenvalues. If z is an eige vector of J, then $y := Tz$ is an eigenvector of A.

∎

For the proof, we refer to any standard book on linear algebra, for example The columns of T corresponding to defective eigenvalues contain, in additio to the eigenvectors, additional vectors called *principal vectors*.

We are now ready to return to our discussion of the behavior of A^n. This discussion is also interesting in the study of iterative methods.

Theorem 1.1.

a) $A^n \to 0$ as $n \to \infty$ iff all eigenvalues of A lie in the interior of the unit disk.

b) A^n is bounded as $n \to \infty$ iff all eigenvalues of A lie in or on the boundary of the unit disk with no defective eigenvalues on the boundary.

∎

Proof. The theorem can be proved in a manner analogous to our handling of the diagonalizable case, but for variety we shall take a slightly different approach.

By induction we see that $A^n = TJ^nT^{-1}$ where J^n is block diagonal with blocks

$$(1.4) \qquad J_i^n = \left(\lambda_i I_{q_i} + S_{q_i}\right)^n = \sum_{j=0}^{q_i-1} \binom{n}{j} \lambda_i^{n-j} S_{q_i}^j \qquad (i = 1,2,\ldots,p)$$

where $\binom{n}{j} = 0$ if $j > n$. Note that the binomial expansion for matrices as used here is only valid since the matrices involved commute: $I_{q_i} S_{q_i} = S_{q_i} I_{q_i}$. Here the summation involves at most q_i terms since $S_{q_i}^{q_i} = 0$. If $|\lambda_i| < 1$, then $J_i^n \to 0$ as $n \to 0$ as $n \to \infty$. If $|\lambda_i| > 1$, then $J_i^n \to \infty$ since the diagonal element are λ_i^n. For $|\lambda_i| = 1$, we have in the non-defective case that $J_i^n = \lambda_i^n I$, which is bounded, but in the defective case the element in position $(1, q_i)$, for example, is $\binom{n}{q_i - 1} \lambda_i^{n-q_i+1}$, which is unbounded. This proves the theorem. $\blacksquare$

As a by-product of the proof, particularly equation (1.4), we obtain the following result:

Theorem 1.2. The matrix A^n can be written as a sum of terms of the form $P_i(n)\lambda_i^n$, where λ_i is an eigenvalue of A with a corresponding Jordan block of order q_i and $P_i(n)$ is a polynomial of degree $q_i - 1$ with uniquely determined matrix coefficients.

2. Difference equations of higher order and associated matrices

Consider now a homogeneous, scalar difference equation of k:th order,

$$\sum_{j=0}^{k} \alpha_j y_{n+j} = 0, \quad (\alpha_k \neq 0), \tag{2.1}$$

with characteristic polynomial

$$\varphi(\lambda) = \sum_{j=0}^{k} \alpha_j \lambda^j. \tag{2.2}$$

The equation (2.1) can be written in operator form as $\varphi(E)y_n = 0$, where E is the translation operator defined by $Ey_n := y_{n+1}$. We can also write (2.1) in matrix-vector form as

$$Y_{n+1} = AY_n \tag{2.3}$$

where

$$Y_j = \begin{bmatrix} y_j \\ y_{j+1} \\ \vdots \\ y_{j+k-1} \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ \frac{-\alpha_0}{\alpha_k} & \frac{-\alpha_1}{\alpha_k} & \frac{-\alpha_2}{\alpha_k} & \dots & \frac{-\alpha_{k-1}}{\alpha_k} \end{bmatrix}. \tag{2.4}$$

We call A the *companion matrix* (for the polynomial φ).

Theorem 2.1. The characteristic polynomial of the *companion matrix* given in (2.4) ~~has~~ is proportional to $\varphi(\lambda)$. ~~as its characteristic polynomial~~. The eigenvalue λ_i possesses only the single eigenvector $[1, \lambda_i, \lambda_i^2, \dots, \lambda_i^{k-1}]^T$. (Multiple eigenvalues are hence always defective, and each eigenvalue appears in only one Jordan block).

Proof. Suppose $Av = \lambda v$ with $v = (v_1, v_2, \dots, v_n)^T$. Then

$$\begin{aligned} v_2 &= \lambda v_1 \\ v_3 &= \lambda v_2 = \lambda^2 v_1 \\ &\dots \\ v_k &= \lambda v_{k-1} = \lambda^{k-1} v_1 \\ -\frac{1}{\alpha_k} \sum_{j=0}^{k-1} \alpha_j v_{j+1} &= \lambda v_k \iff \sum_{j=0}^{k} \alpha_j \lambda^j v_1 = 0. \end{aligned}$$

Hence the only eigenvector is

$$v = v_1(1, \lambda, \lambda^2, \dots, \lambda^{k-1})^T.$$

∎

Definition. A polynomial $\varphi(z)$ is said to satisfy the *root condition* if its roots lie in the closed unit disk with only simple zeros on the boundary.

Corollary to Theorems 1.1 and 2.1

The solution of the difference equation (2.1) is bounded as $n \to \infty$ for every choice of initial values $y_0, y_1, \ldots, y_{k-1}$ if and only if the characteristic polynomial (2.2) satisfies the root condition. ∎

From theorems 1.2 and 2.1 we obtain the well-known result that *the general solution of the scalar difference equation* $\varphi(E)y_n = 0$ (if $\varphi(0) \neq 0$) *is a sum of terms of the form* $\psi_i(n)\lambda_i^n$ *where* λ_i *is a root of* $\varphi(\lambda)$ *of multiplicity* m_i, *and* ψ_i *is an arbitrary polynomial of degree* $m_i - 1$ which is determined by the initial conditions. We omit the details since a more direct proof can be found in Dahlquist & Björck [2]. (Note that the polynomial P_i with matrix coefficients which appeared in Theorem 1.2 was uniquely determined.)

Example 2.1. Consider the difference equation

$$y_{n+1} - 2y_n + (1+\mu^2)y_{n-1} = 0; \quad y_0 = a, \quad y_1 = a + b.$$

The characteristic equation is $(\lambda - 1)^2 + \mu^2 = 0$ with solutions $\lambda_1 = 1 + i\mu$, $\lambda_2 = 1 - i\mu$. The difference equation thus has unbounded solutions for any μ, since the root condition is never satisfied.

If $\mu \neq 0$. The general solution is $y_n = p(1+i\mu)^n + g(1-i\mu)^n$. The initial conditions give

$$y_n = \frac{1}{2}\left(a - \frac{ib}{\mu}\right)(1+i\mu)^n + \frac{1}{2}\left(a + \frac{ib}{\mu}\right)(1-i\mu)^n \tag{2.5}$$

If $\mu = 0$. The general solution is $y_n = p + qn$ and the initial conditions give $y_n = a + bn$. The solution given by (2.5) is not in a suitable form when $|\mu| \ll 1$. Another form is obtained in the following way. Set

$$1 \pm i\mu = (1+\mu^2)^{1/2} e^{\pm i\alpha}, \qquad \alpha = \arctan \mu.$$

So

$$\begin{aligned} y_n &= \left(1+\mu^2\right)^{n/2}\left[\frac{1}{2}\left(a - \frac{ib}{\mu}\right)e^{in\alpha} + \frac{1}{2}\left(a + \frac{ib}{\mu}\right)e^{-in\alpha}\right] \\ &= \left(1+\mu^2\right)^{n/2}\left[a\cos n\alpha + b\,\frac{\sin n\alpha}{\mu}\right] \end{aligned} \tag{2.6}$$

Since $\alpha/\mu \to 1$ as $\mu \to 0$, we see that $y_n \to a + bn$ as $\mu \to 0$. The form (2.6) is relatively well conditioned for $|\mu| \ll 1$. ∎

Companion matrices arise naturally when studying higher order difference equations with specified initial values. Another important class of matrices come from difference equations with specified boundary values. Such problems often lead to band matrices of a special form. One particular application is to the so-called *method of lines* approach for solving partial differentia equations. In this method one discretizes the space variable(s) while maintaining continuity in time. The partial differential equation is thus approximated by a system of ordinary differential equations.

Example 2.2. The heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \qquad 0 \le x \le 1, \quad t > 0$$

with boundary conditions $u(0,t) = u(1,t) = 0$ and initial conditions $u(x,0) = f(x)$. We discretize the interval (0.1) by taking

$$\Delta x = 1/N, \; x_i := i\Delta x \text{ for } i = 0,1,\dots,N.$$

Using central-difference approximations to the space derivatives and defining the new variables $u_i(t)$, intended to be approximations to $u(x_i,t)$, we obtain

$$\text{(2.7)} \qquad \frac{du_i}{dt} = \frac{u_{i-1} - 2u_i + u_{i+1}}{(\Delta x)^2} \qquad i = 1,2,\dots,N-1$$

$$u_0 = u_N = 0$$

Let $u = (u_1,u_2,\dots,u_{N-1})^T$. Then the system (2.7) can be written as

$$\text{(2.8)} \qquad \frac{du}{dt} = Au$$

where

$$\text{(2.9)} \qquad A = \begin{bmatrix} b & c & & & & O \\ a & b & c & & & \\ & a & b & c & & \\ & & \ddots & \ddots & \ddots & \\ & & & a & b & c \\ O & & & & a & b \end{bmatrix}$$

with $a = c = N^2$, $b = -2N^2$. ∎

We will return to this example after deriving expressions for the eigenvalues and eigenvectors of the general matrix given in (2.9). To this end we first consider the special case $a = c = 1$, $b = 0$.

Example. Solve the eigenvalue problem

$$y_{n+1} - \lambda y_n + y_{n-1} = 0 \tag{2.10}$$

with boundary conditions

$$y_0 = y_N = 0.$$

The characteristic equation can be written as

$$u^2 - \lambda u + 1 = 0 \tag{2.11}$$

with roots u_1 and $u_2 = u_1^{-1}$. Dropping the index on u_1, the general solution of (2.10) can be written

$$y_n = \alpha u^n + \beta u^{-n}.$$

The boundary conditions give

$$\left.\begin{array}{l} \alpha + \beta = 0 \\ \alpha u^N + \beta u^{-N} = 0 \end{array}\right\} \Rightarrow \alpha(u^N - u^{-N}) = 0$$

and so

$$u^{2N} = 1, \quad u = \exp\left(\frac{\pi i j}{N}\right), \quad j = 0,1,\ldots,2N-1$$

$$y_n = 2\alpha i \sin\frac{\pi j n}{N}, \quad j = 1,2,\ldots,N-1.$$

(The remaining values of j give either the null solution or a repetition of those determined by $1 \le j < N$). From (2.10) and (2.11) it follows that

$$\lambda = u + 1/u = 2\cos\frac{\pi j}{N} \quad j = 1,2,\ldots,N-1.$$

∎

From the computations of this example we obtain the general result.

Theorem 2.2. The tridiagonal $(N-1)\times(N-1)$ matrix A given in (2.9) has eigenvalues

$$\lambda_j = b + 2c\sqrt{a/c}\cos\frac{\pi j}{N} \qquad (j = 1,2,\ldots,N-1)$$

The corresponding eigenvectors have as their n:th component

$$y_n = (\sqrt{a/c})^n \sin\frac{\pi j n}{N} \qquad (j = 1,2,\ldots,N-1).$$

Proof. We attempt to "symmetrize" the matrix by means of a scaling transformation $D = \mathrm{diag}(d_i)$, $d_0 = 1$. The elements of $D^{-1}AD$ are given by

$$(D^{-1}AD)_{ij} = \begin{cases} cd_j/d_i & \text{if } j = i+1 \\ b & \text{if } j = i \\ ad_j/d_i & \text{if } j = i-1 \end{cases}$$

Thus the resulting matrix is symmetric if

$$cd_j/d_i = ad_i/d_j \quad (j = i+1)$$

i.e. if $d_{i+1}/d_i = \sqrt{a/c}$. Set $d_i := (\sqrt{a/c})^i$ and then the off-diagonal elements become $c\sqrt{a/c} = \sqrt{ac}$. Thus

$$D^{-1}AD = bI + \sqrt{ac}\,A_1 \tag{2.12}$$

where A_1 is given by the matrix in (2.9) with $b = 0$, $a = c = 1$. The eigenvalue problem $A_1 y = \lambda y$ gives

$$\begin{cases} y_{n-1} + y_{n+1} = \lambda y_n \\ y_0 = y_N = 0 \end{cases}$$

i.e. the problem of the previous example. Thus

$$\lambda = 2\cos\frac{\pi j}{N}; \quad y_n = \sin\frac{\pi jn}{N}.$$

A has the same eigenvalues as $D^{-1}AD$, which according to (2.12) are simply,

$$\lambda = b + 2\sqrt{ac}\cos\frac{\pi j}{N}.$$

If $D^{-1}ADy = \lambda y$ then $A(Dy) = \lambda(Dy)$. The eigenvectors thus have components $Dy_n = (\sqrt{a/c})^n \sin\frac{\pi jn}{N}$. (We need not worry about the sign of the square root. We shall interpret $\sqrt{ac}$ as $c\sqrt{a/c}$).

■

Returning now to example 2.2, we see that the matrix A occuring in the system (2.8) has eigenvalues $-2N^2 + 2N^2\cos(\pi j/N) = -4N^2\sin^2(\pi j/2N)$. Notice the great spread of eigenvalues from approximately $-4N^2$ to approximately $-\pi^2$. The larger N is, the more "stiff" the system becomes. Notice also that the transformation of A to diagonal form is equivalent to a discrete Fourier transform (sine-transform), and that the j'th eigenvalue tends to $-\pi^2 j^2$ when $N\to\infty$, which is the j'th eigenvalue of the differential operator d^2/dx^2 with boundary conditions, $u(0) = u(1) = 0$.

3. Matrix norms and logarithmic norms

Many important problems concerning difference and differential equations can be handled in a particularly elegant manner through the use of *norms*. In this section we review the most important vector and matrix norms and simultaneously develop the so-called *logarithmic norms*.

A *vector norm* is a real-valued function $\|y\|$ (defined for all y in some vector space V), which satisfies the following axioms

1) $\|y\| > 0$ for $y \neq 0$
2) $\|\alpha y\| = |\alpha| \cdot \|y\|$ $(\alpha \in C)$
3) $\|y_1 + y_2\| \leq \|y_1\| + \|y_2\|$ (triangle inequality)

From 3) we obtain

3*) $\|y_1 - y_2\| \geq \|y_1\| - \|y_2\|$.

From 2) and 3) it follows that $\|y\|$ is a *convex function* of y, since for $0 \leq t \leq 1$ we have

$$\|(1-t)y_0 + ty_1\| \leq (1-t)\,\|y_0\| + t\,\|y_1\| \;.$$

It then follows that the *unit ball* $\{y : \|y\| \leq 1\}$ is a convex set which is symmetric about the origin (since $\|-y\| = \|y\|$). We note, in passing, that conversely any convex, symmetric set with a non-empty interior, can be taken as a unit ball to define a norm, the value of which at an arbitrary point y, is obtained by axiom 2, choosing α so that αy lies on the boundary of S.

The triangle inequality easily generalizes to sums with an arbitrary number of terms and in the limit even to infinite series and integrals. For example,

$$\Big\|\sum_{n=0}^{\infty} \alpha_n y_n\Big\| \leq \sum_{n=0}^{\infty} |\alpha_n|\,\|y_n\|,$$

$$\Big\|\int_0^1 \alpha(t)y(t)dt\Big\| \leq \int_0^1 |\alpha(t)|\cdot\|y(t)\|dt \leq \max_{0\leq t\leq 1}\|y(t)\|\cdot\int_0^1|\alpha(t)|dt.$$

To a given vector norm $\|\cdot\|$ there corresponds a *subordinate matrix norm* defined by

$$\|A\| = \max_{\substack{x\in V\\ x\neq 0}} \frac{\|Ax\|}{\|x\|}\,.$$

It follows easily from the definition that *matrix norms also satisfy axiom* 1) – 3) and their consequences, e.g. convexity. We also have the following *submultiplicative* properties:

4) $\|Ax\| \le \|A\| \cdot \|x\|$

5) $\|AB\| \le \|A\| \cdot \|B\|$.

Here A and B can be rectangular matrices.

There is a more general axiomatic definition of matrix norms, but the present one suffices for our needs.

To a given vector norm $\|\cdot\|$ there also corresponds a *logarithmic norm* $\mu(A)$, defined by

$$(3.1) \qquad \mu(A) = \lim_{h\downarrow 0} \frac{\|I+hA\|-1}{h} .$$

This limit exists for every norm $\|\cdot\|$ and every matrix A. This follows from the convexity of the norm, which guarantees that the difference ratio $\frac{\|I+hA\|-1}{h}$ decreases monotonically when h decreases, and from the triangle inequality 3*), which bounds this ratio below by $\frac{\|I\|-h\|A\|-1}{h} = -\|A\|$.

The following *properties* follow directly from the definition:

(3.2a) $\mu(\alpha A) = \alpha\mu(A)$ *only* if $\alpha \ge 0$

(3.2b) $\mu(A_1+A_2) \le \mu(A_1) + \mu(A_2)$ (subadditivity)

(3.2c) $-\|A\| \le \mu(A) \le \|A\|$.

It follows that $\mu(A)$ *is a convex function*, for take $0 \le t \le 1$. Then

$$\mu((1-t)A_0 + tA_1) \le \mu((1-t)A_0) + \mu(tA_1) = (1-t)\mu(A_0) + t\mu(A_1).$$

For the special case $A = zI$, $z \in C$ we obtain

$$\|A\| = \|zI\| = |z| \cdot \|I\| = |z|$$

$$\mu(A) = \lim_{h\downarrow 0} \frac{\|I+hzI\|-1}{h} = \lim_{h\downarrow 0} \frac{|1+hz|-1}{h} = \mathrm{Re}\, z .$$

(Verify as an exercise the last equality)

Observe that $\mu(A)$ can be negative. Just as $\|A\|$ can be viewed as a generalization of the absolute value, so can $\mu(A)$ be viewed as a generalization of the real part of a complex number. Additional interesting general properties will be found in the exercises.

Definition. The set of eigenvalues of A is called the *spectrum* of A. With A we can associate the following two real numbers:

$$\rho(A) := \max_{\lambda \in \text{spectrum}(A)} |\lambda|, \quad \text{the } \textit{spectral radius} \text{ of } A,$$

$$\alpha(A) := \max_{\lambda \in \text{spectrum}(A)} \operatorname{Re} \lambda, \quad \text{the } \textit{spectral abscissa} \text{ of } A.$$

The most commonly used norms are the ℓ_1-, ℓ_2- and ℓ_∞-norms and transformations of these. The following table gives expressions for these norms. Here $x = [x_i]$, $A = [a_{ij}]$, and $\sum_j{}'$ means $\sum_{\substack{j \\ j \neq i}}$.

p	Vector norm $\|x\|_p$	Matrix norm $\|A\|_p$	Logarithmic norm $\mu_p(A)$
∞	$\|x\|_\infty = \max_i \|x_i\|$	$\|A\|_\infty = \max_i \sum_j \|a_{ij}\|$	$\mu_\infty(A) = \max_i \left[\operatorname{Re} a_{ii} + \sum_j{}' \|a_{ij}\|\right]$
1	$\|x\|_1 = \sum_i \|x_i\|$	$\|A\|_1 = \max_j \sum_i \|a_{ij}\|$	$\mu_1(A) = \max_j \left[\operatorname{Re} a_{jj} + \sum_i{}' \|a_{ij}\|\right]$
2	$\|x\|_2 = \left[\sum_i \|x_i\|^2\right]^{1/2}$	$\|A\|_2 = [\rho(A^H A)]^{1/2}$	$\mu_2(A) = \alpha\left(\frac{A+A^H}{2}\right)$

Note that a *row* vector y^T is to be considered as a $1 \times n$ matrix, and that in general $\|y^T\| \neq \|y\|$. (The ℓ_2-norm is an exception in this respect.)

We give two examples of verifying the results in the table:

1) Assume the expression for $\|A\|_\infty$ is known. It follows that

$$\mu_\infty(A) = \lim_{h \downarrow 0} \max_i \left[\frac{|1+ha_{ii}| + h\sum_j{}' |a_{ij}| - 1}{h}\right]$$

$$= \max_i \left[\lim_{h \downarrow 0}\left(\frac{|1+ha_{ii}| - 1}{h}\right) + \sum_j{}' |a_{ij}|\right]$$

$$= \max_i \left[\operatorname{Re} a_{ii} + \sum_j{}' |a_{ij}|\right].$$

2) Assume the expression for $\|A\|_2$ is known. Let μ_i be an eigenvalue of the Hermitian matrix $\frac{A^H+A}{2}$. Note that μ_i is real. Let $\|\cdot\| := \|\cdot\|_2$ to simplify the notation. Then

$$\begin{aligned}
\|I + hA\|^2 &= \rho((I + hA^H)(I + hA)) \\
&= \rho(I + h(A^H + A) + h^2 A^H A) \\
&= \|I + h(A^H + A) + h^2 A^H A\| \quad \text{since } \rho(B) = \|B\|_2 \text{ for a Hermitian matrix } B \\
&= \|I + h(A^H + A)\| + O(h^2) \\
&= \max_i (1 + 2h\mu_i) + O(h^2)
\end{aligned}$$

and thus

$$\frac{\|I + hA\| - 1}{h} = \frac{\|I + hA\|^2 - 1}{h(\|I + hA\| + 1)}$$
$$= \max_i \frac{(1+2h\mu_i) - 1 + O(h^2)}{h(1+O(h))}$$
$$= \max_i \frac{2\mu_i + O(h)}{2 + O(h)}$$

In the limit as $h \downarrow 0$ we obtain

$$\mu_2(A) = \max_i \mu_i$$
$$= \alpha\left(\frac{A+A^H}{2}\right)$$

∎

Theorem 3.1. For every matrix or logarithmic norm subordinate to a vector norm,

$$\|A\| \geq \rho(A) \text{ and } \mu(A) \geq \alpha(A).$$

Note. a) The quantities $\rho(A)$ and $\alpha(A)$ are independent of the norm used here

b) In the ℓ_∞- and ℓ_1-case we obtain from this bounds for the spectrum which can also be obtained by the Gerschgorin theorem.

c) According to the table, we have equality in Theorem 3.1 in certain cases. For example, in the ℓ_2-case when A is Hermitian and for every ℓ_p-norm when A is diagonal matrix.

Proof. By the definition of the spectral radius, $\exists\, \hat{\lambda}, \hat{x}$ with $|\hat{\lambda}| = \rho(A)$ such that $A\hat{x} = \hat{\lambda}\hat{x}$. Thus,

$$\|A\hat{x}\| = \rho(A)\ \|\hat{x}\|$$

and so

$$\|A\| = \max_x \frac{\|Ax\|}{\|x\|} \geq \frac{\|A\hat{x}\|}{\|\hat{x}\|} = \rho(A).$$

There also exists $\tilde{\lambda}$, $\tilde{x}$, independent of h, such that $A\tilde{x} = \lambda\tilde{x}$ and $\mathrm{Re}\,\tilde{\lambda} = \alpha(A)$ Thus,

$$\|I + hA\| \cdot \|\tilde{x}\| \geq \|(I + hA)\tilde{x}\| = |1 + h\tilde{\lambda}| \cdot \|\tilde{x}\|$$

and so

$$\mu(A) = \lim_{h \downarrow 0} \frac{\|I + hA\| - 1}{h} \geq \lim_{h \downarrow 0} \frac{|1 + h\tilde{\lambda}| - 1}{h} = \mathrm{Re}\,\tilde{\lambda} = \alpha(A).$$

∎

Although the ℓ_2-case is easy to work with in many ways, the matrix and logarithmic norms are often difficult to compute. From Theorem 3.1 we obtain the following useful inequalities:

$$(3.3) \qquad \|A\|_2^2 = \rho(A^H A) \le \|A^H A\|_\infty \le \|A^H\|_\infty \cdot \|A\|_\infty = \|A\|_1 \cdot \|A\|_\infty$$
$$\mu_2(A) = \alpha\left(\frac{A+A^H}{2}\right) \le \mu_\infty\left(\frac{A+A^H}{2}\right).$$

Many results for the ℓ_2-norm can be generalized to arbitrary innerproduct norms.

Definition. An *innerproduct* (x,y) in a complex vector space satisfies the following axioms:

a) $(x,y) = \overline{(y,x)}$, symmetry (the bar indicates conjugation)

b) $(\alpha x, \beta y) = \bar{\alpha}\beta(x,y)$ } linearity

c) $(x, y+z) = (x,y) + (x,z)$ }

d) $(x,x) > 0$ for $x \neq 0$, positivity.

It is easy to show that $\|x\| := \sqrt{(x,x)}$ satisfies the axioms of a vector norm. With this norm we also obtain *Schwarz' inequality:*

$$(3.4) \qquad |(x,y)| \le \|x\| \cdot \|y\|.$$

We can easily obtain an expression for the logarithmic norm subordinate to this vector norm. From the axioms for an inner product we find that

$$\frac{\|x+hAx\| - \|x\|}{h\|x\|} = \frac{\|x+hAx\|^2 - \|x\|^2}{h\|x\|(\|x+hAx\| + \|x\|)}$$
$$= \frac{2h\mathrm{Re}(x,Ax) + h^2\|Ax\|^2}{h\|x\|(\|x+hAx\| + \|x\|)},$$

(verify the last equality), and so

$$\frac{\|I+hA\| - 1}{h} = \max_{x \neq 0} \frac{\mathrm{Re}(x,Ax)}{\|x\|^2} + O(h).$$

Thus, as $h \downarrow 0$, we obtain by the linearity of the inner product,

$$(3.5) \qquad \mu(A) = \max_{\|x\|=1} \mathrm{Re}\,(x,Ax).$$

In $\mathbb{C}^s$ the most general form of an inner product is

$$(3.6) \qquad (x,y) = x^H G y$$

where G is some positive definite Hermitian matrix. The corresponding norm is defined by

$$(3.6') \qquad \|x\|^2 := x^H G x .$$

Note that for $G = I$ this is simply the ℓ_2-norm. Since G is positive definite, the same holds for G^{-1}. Let the Cholesky factorization of G^{-1} be

$$G^{-1} =: TT^H .$$

We can then look at the inner product (3.6) in a different way. Consider the coordinate transformation $x = T\xi$, $y = T\eta$. We can think of ξ and η as the coordinate vectors for x and y when the columns of T are used as a bases. Then, by (3.6),

$$(x,y) = \xi^H T^H (TT^H)^{-1} T\eta = \xi^H \eta$$

i.e. the inner product (3.6) is the usual inner product of the vectors ξ and η. Furthermore $(x,x) = \|\xi\|_2^2 = \|T_x^{-1}\|_2^2$ which relates the inner product norm (3.6') to the usual ℓ_2-norm.

This relation suggests the following generalization of the inner product norm (3.6'). Let $\|\cdot\|$ be *any vector norm* and T be *any nonsingular matrix*. We then define the *T-norm* of x by

$$(3.7) \qquad \|x\|_T := \|T^{-1}x\| .$$

It is easy to verify that $\|\cdot\|_T$ satisfies the norm axioms. The corresponding matrix norm is

$$(3.8) \qquad \|A\|_T = \max_{x \in V} \frac{\|Ax\|_T}{\|x\|_T} = \max_{\xi \in V} \frac{\|T^{-1}AT\xi\|}{\|\xi\|} = \|T^{-1}AT\| .$$

Similarly,

$$(3.9) \qquad \mu_T(A) = \mu(T^{-1}AT).$$

As an application of (3.8) and (3.9), *consider again the important special case of the inner product norm* (3.6'), which we can now write as

$$\|x\|_T^2 = \|T^{-1}x\|_2 = x^H G x$$

where $G = T^{-H}T^{-1}$. Using the table result for the matrix ℓ_2-norm, we have

$$\begin{aligned} \|A\|_T^2 &= \|T^{-1}AT\|_2^2 \\ &= \rho(T^H A^H T^{-H} T^{-1} AT). \end{aligned}$$

Let λ be an eigenvalue of $T^H A^H T^{-H} T^{-1} AT$. Then, since T is nonsingular, we obtain

$$\begin{aligned} 0 &= \det(T^H A^H T^{-H} T^{-1} AT - \lambda I) \\ &= \det(A^H T^{-H} T^{-1} A - \lambda T^{-H} T^{-1}) \\ &= \det(A^H GA - \lambda G), \end{aligned}$$

and hence the matrix norm corresponding to (3.6') is given by the solution of a *generalized eigenvalue problem*

$$\|A\|_T^2 = \max\{\lambda : \det(A^H GA - \lambda G) = 0\}. \tag{3.10}$$

Similarly, for the logarithmic norm we have,

$$\begin{aligned} \mu_T(A) &= \mu_2(T^{-1}AT) \\ &= \alpha(\tfrac{1}{2}(T^{-1}AT + T^H A^H T^{-H})) \\ &= \max\{\lambda : \det(\tfrac{1}{2}(GA + A^H G) - \lambda G) = 0\}. \end{aligned} \tag{3.11}$$

The expressions (3.10) and (3.11) can also be derived in a different manner. From the definition of the matrix norm we have that

$$\|A\|_T^2 = \max_{x^H Gx = 1} x^H A^H GAx.$$

Solving this constrained optimization problem by Lagrange multipliers, for example, again leads to (3.10). See Bellman [3] for more details on this approach. Similarly, for the logarithmic norm, (3.5) gives

$$\begin{aligned} \mu_T(A) &= \max_{x^H Gx = 1} \operatorname{Re} x^H GAx \\ &= \max_{x^H Gx = 1} \tfrac{1}{2} x^H (GA + A^H G)x \end{aligned}$$

which again leads to (3.11).

Returning now to the general T-norm corresponding to an *arbitrary* norm $\|\cdot\|$, we see from (3.8) that $\|A\|_T \le \|T^{-1}\| \cdot \|A\| \cdot \|T\|$ and that, if $T^{-1}AT = B$, $\|B\| = \|TBT^{-1}\| \le \|T^{-1}\|_T \cdot \|T\|_T \cdot \|B\|_T$. Observe also that $\|T\|_T = \|T^{-1}TT\| = \|T\|$ and similarly $\|T^{-1}\|_T = \|T^{-1}\|$. From this we obtain the important *norm equivalence inequality*, where $\operatorname{cond}(T) = \|T^{-1}\| \cdot \|T\|$ is the *condition number* of T.

$$\frac{\|A\|}{\operatorname{cond}(T)} \le \|A\|_T \le \|A\| \cdot \operatorname{cond}(T). \tag{3.12}$$

Similar inequalities do *not* hold for the logarithmic norms. For the vector norm we have

$$\frac{\|x\|}{\|T\|} \le \|x\|_T \le \|T^{-1}\| \cdot \|x\|. \tag{3.13}$$

The following important theorem provides a sort of converse to Theorem 3.1.

Theorem 3.2.

For any ℓ_p-norm $\|\cdot\|$ the following hold:

a) If A has no defective eigenvalues with absolute value $\rho(A)$ then $\exists T$ such that

$$\|A\|_T = \rho(A).$$

b) If A has defective eigenvalues with absolute value $\rho(A)$, then for every $\varepsilon > 0$ there exists a matrix $T(\varepsilon)$ such that

$$\|A\|_{T(\varepsilon)} < \rho(A) + \varepsilon.$$

In this case, $\lim_{\varepsilon\to 0} \|T(\varepsilon)\| \cdot \|T^{-1}(\varepsilon)\| = \infty$.

c) If A has no defective eigenvalue with real part $\alpha(A)$ then there exists a matrix T such that

$$\mu_T(A) = \alpha(A).$$

d) If A has defective eigenvalues with real part $\alpha(A)$, then for any $\varepsilon > 0$ there is a matrix $T(\varepsilon)$ such that

$$\mu_{T(\varepsilon)}(A) < \alpha(A) + \varepsilon.$$

Note. In general the same matrix T will not satisfy both a) and c), nor will the same $T(\varepsilon)$ satisfy both b) and d).

Proof. We show only a) and b). The proof for c) and d) is similar. Set $\rho := \rho(A)$. Since $\|A\|_T = \|T^{-1}AT\|$, our goal will be to use a similarity transformation to transform A to a matrix $T^{-1}AT$ for which it is easy to show that the norm is sufficiently close to the spectral radius. If A is diagonalizable, we can simply take T as the diagonalizing transformation. Then clearly $\|A\|_T = \|T^{-1}AT\| = \max_i |\lambda_i| = \rho$. For the proof of the theorem we reduce A to a matrix which is sufficiently close to being diagonal. To be exact, we use a variation of the Jordan Form in which the nonzero superdiagonal elements have arbitrary small values. To achieve this, we first let $\hat{J} = \hat{T}^{-1}A\hat{T}$ be the Jordan Canonical Form of A as in Lemma 1.1. So $\hat{J} = \text{blockdiag}(\hat{J}_i)$ with $\hat{J}_i = \lambda I + S_i$, a Jordan block of order q_i. Here S_i is a shift matrix of order q_i. The proof is based on the following observations (also useful in other contexts):

(i) For $\delta > 0$ let $D_i(S) := \text{diag}(1, \delta, \ldots, \delta^{q_i - 1})$ and let $J_i := D_i^{-1}(\delta)\hat{J}_i D_i(\delta)$. Then $J_i = \lambda_i I + \delta S_i$. Verify this!

(ii) If $J = \text{blockdiag}(T_i)$ then $\|J\| = \max_i \|J_i\|$ for any ℓ_p-norm.

(iii) $\|S_i\| \le 1$ in any ℓ_p-norm for a shift matrix S_i. This is because $S_i x = (x_2, x_3, \ldots, x_{q_i}, 0)^T$ so $\|S_i x\| \le \|x\|$ for all vectors x. From this it follows that $\|\lambda_i I + \delta S_i\| \le |\lambda_i| + \delta$.

To simplify the discussion below, let $R := \{i : |\lambda_i| = \rho\}$, the set of indices of the maximal eigenvalues.

For the proof of a), we observe that by assumption if $i \in R$ then $q_i = 1$ and so $J_i = \lambda_i$, $D_i(\delta) \equiv 1$. Hence

$$\|J_i\| = \rho \quad \text{for} \quad i \in R.$$

For $i \notin R$, let $\delta_i = \rho - |\lambda_i| > 1$. Now take $D := \text{blockdiag}(D_i(\delta_i))$, $T := \hat{T}D$, $J := T^{-1}AT$. Then $J_i = \lambda_i I + \delta_i S_i$ and so

$$\begin{aligned} \|J_i\| &\le |\lambda_i| + \delta_i \\ &\le \rho \quad \text{for} \quad i \notin R, \end{aligned}$$

by our choice of δ_i and (iii). It follows from (ii) that

$$\|A\|_T = \|J\| = \max_i \|J_i\| = \rho,$$

which completes the proof of a).

For the proof of b), let $\varepsilon > 0$ be given. We now choose

$$\delta_i = \begin{cases} \varepsilon \text{ if } i \in R \\ \rho - |\lambda_i| \text{ if } i \notin R. \end{cases}$$

Take $D(\varepsilon) := \text{blockdiag}(D_i(\delta_i))$, $T(\varepsilon) := TD(\varepsilon)$, $J(\varepsilon) := T^{-1}(\varepsilon)AT(\varepsilon)$. Now for $i \in R$ we have $\|J_i(\varepsilon)\| \le |\lambda_i| + \delta_i = \rho + \varepsilon$ and for $i \notin R$. $\|J_i(\varepsilon)\| \le \rho$ as before. Hence

$$\|A\|_{T(\varepsilon)} = \max_i \|J_i(\varepsilon)\| \le \rho + \varepsilon.$$

This completes the proof of part b). Note that in this case $\|D_i(\varepsilon)\| \cdot \|D_i^{-1}(\varepsilon)\| = \varepsilon^{1-q_i}$ for $i \in R$ and so, for ε sufficiently small. $\|D(\varepsilon)\| \cdot \|D^{-1}(\varepsilon)\| = \max_i \|D_i(\varepsilon)\| \cdot \|D_i^{-1}(\varepsilon)\| = \varepsilon^{1-q^*}$ where $q^* = \max_{i \in R} q_i$. We thus have

$$\|T(\varepsilon)\| \cdot \|T^{-1}(\varepsilon)\| \le \|\hat{T}\| \cdot \|\hat{T}^{-1}\| \cdot \|D(\varepsilon)\| \cdot \|D^{-1}(\varepsilon)\| = O(\varepsilon^{1-q^*})$$

Furthermore,

$$\|T(\varepsilon)\| \cdot \|T^{-1}(\varepsilon)\| \ge \frac{\|D(\varepsilon)\| \cdot \|D^{-1}(\varepsilon)\|}{\|\hat{T}\| \cdot \|\hat{T}^{-1}\|}$$

so the condition number of T increases rapidly as $\varepsilon \to 0$.

■

In Theorem 3.2 the T-norm depends on the matrix A. It is not, however uniquely determined. If the elements of A are analytic functions of a complex variable q, then T and T(ε) can be chosen as continuous function of q. Unfortunately, the construction used in the proof does not provide this in general, since the Jordan Form can be a discontinuous function of q.

Exercises.

3.1) Verify the results of the table on p.3.

3.2a) Show that $\|A^{-1}\| \leq 1/|\mu(A)|$ if $\mu(A) < 0$.

b) Show that if $a_{ii} > \sum_{j \neq i} |a_{ij}| \; \forall i$ then

$$\|A^{-1}\|_\infty \leq \frac{1}{\min_i (a_{ii} - \sum_{j \neq i} |a_{ij}|)} .$$

3.3) Consider two norms $|x|$, $\|x\|$ and the corresponding matrix norms. Assume that for all x

$$\alpha|x| \leq \|x\| \leq \beta \cdot |x| .$$

Is (3.12) a particular case of this?

4. The use of norms in difference and differential equations

We are now ready to return to the study of the boundedness of solutions of difference and differential equations. Recall our original example $y_{n+1} = Ay_n$. If T is a nonsingular matrix, then (3.12) yields

$$\|A^n\| \le \operatorname{cond}(T)\cdot\|A^n\|_T \le \operatorname{cond}(T)\cdot\|A\|_T^n.$$

Thus $A^n \to 0$ as $n\to\infty$ if we can find a T such that $\|A\|_T < 1$ and A^n at least remains bounded if $\|A\|_T \le 1$. In view of this we see that Theorem 1.1 follows directly from theorems 3.1 and 3.2. For if $\rho(A) < 1$ we can always find a T such that $\|A\|_T < \rho(A) + \varepsilon < 1$ and if $\rho(A) \le 1$ with no defective eigenvalues of modulus 1 then we can find a T such that $\|A\|_T = \rho(A) \le 1$. This gives an idea of the use of matrix norms in difference equations.

For differential equations the logarithmic norms play a similar role. Consider the linear system

$$y'(t) = A(t)y(t)\ . \tag{4.1}$$

We will see that $\mu(A(t)) < 0\ \forall t$ *is a sufficient condition for* $y(t)$ *to be bounded as* $t\to\infty$.

Let $y(t)$ be a solution of (4.1) and assume that $\|A'(t)\|$ is bounded for $t \in [a,b]$. By Taylor's theorem,

$$\begin{aligned} y(t+h) &= y(t) + hy'(t) + O(h^2) \\ &= (I + hA(t))y(t) + O(h^2). \end{aligned}$$

So

$$\|y(t+h)\| \le \|I + hA(t)\|\cdot\|y(t)\| + O(h^2)$$

$$\frac{\|y(t+h)\| - \|y(t)\|}{h\|y(t)\|} \le \frac{\|I + hA(t)\| - 1}{h} + O(h).$$

The difference ratio on the right, whose limit defines $\mu(A(t))$, is thus an upper bound on the relative change per time step of $\|y(t)\|$. Now let $h\to 0$ to obtain

$$\frac{d\|y(t)\|/dt}{\|y(t)\|} \le \mu(A(t)) \tag{4.2}$$

and hence by integrating,

$$\log\|y(t)\| - \log\|y(0)\| \le \int_0^t \mu(A(\tau))d\tau \tag{4.3}$$

Thus,

$$(4.4) \qquad \|y(t)\| \le \exp\left(\int_0^t \mu(A(\tau))\,d\tau\right)\|y(0)\|$$

If $\hat{\mu}$ is an upper bound on $\mu(A(t))$ for $t \in [0,\infty)$ then we obtain

$$(4.4') \qquad \|y(t)\| \le e^{\hat{\mu}t}\|y(0)\| \quad \text{for } t \ge 0$$

which proves the assertion, namely that $\hat{\mu} \le 0$ implies $y(t)$ remains bounded as $t \to \infty$.

We could have taken a different approach. Note that

$$\frac{d\|y(t)\|}{dt} \le \left\|\frac{dy}{dt}\right\| \le \|A(t)\| \cdot \|y(t)\|$$

and hence

$$\frac{d\|y(t)\|/dt}{\|y(t)\|} \le \|A(t)\| .$$

By (3.2c), this is never better than (4.4). In general it is much worse. In particular, it does not lead to any conditions for the boundedness of the solutions as $t \to \infty$.

Let us specialize now to the case of a constant matrix A in (4.1). In this case $\mu(A) \le 0$ is a sufficient condition for boundedness. If our underlying vector norm is an *inner-product norm*, then by (3.5) the condition $\mu(A) \le 0$ is equivalent to the condition

$$\mathrm{Re}(x,Ax) \le 0 \ \forall x,$$

a so-called *monotonicity condition* for the matrix A. If the inner product is given by $(x,y) = x^H Gy$, then this means that $GA + A^H G$ should be negative semi-definite, which is the usual condition in the Liapunov stability theory.

The inequality (4.4) holds for any vector norm and its corresponding logarithmic norm. Clearly the solution $y(t)$ of $y'(t) = Ay(t)$ is bounded if we can find some T-norm such that $\mu_T(A) < 0$. By Theorem 3.2, such a T exists provided $\alpha(A) \le 0$ and A has no defective eigenvalues λ with $\mathrm{Re}\,\lambda = 0$. Moreover, $y(t) \to 0$ as $t \to \infty$ if there exists a T-norm such that $\mu_T(A) < 0$, i.e. if $\alpha(A) < 0$.

For the differential equation with constant coefficients, $y' = Ay$, we thus have two types of conditions. One, a sufficient condition, involves the logarithmic norm and the other, a necessary and sufficient condition, involves the spectral abscissa. Similarly, for the difference equation $y_{n+1} = Ay_n$, we obtained conditions involving either the matrix norm or or the spectral radius. The advantage of the norm conditions lies in their applicability to variable coefficient (and even nonlinear) problems. We will see that the spectral conditions do not generalize in the same way. The price we pay for using the norm conditions is that, since they are merely sufficient conditions, the sharpness of our results will depend heavily on the choice of norm. Although Theorem 3.2 gives a hint of how to choose a norm, it must be remembered that the T-norm defined in that theorem depends on the matrix A, while inequalities like (4.2) and (4.4) assume a fixed norm. We therefore need a more general discussion of the use of norms for difference and differential equations with variable coefficients.

Consider the difference equation

$$y_{n+1} = A_n y_n \tag{4.5}$$

with variable coefficients. If $\|A_n\|_T \leq 1$ $\forall n$ for some fixed T-norm, then by the submultiplicativity of the matrix norm we have that y_n is bounded as $n \to \infty$:

$$\begin{aligned} \|y_n\|_T &= \|A_{n-1} A_{n-2}, \ldots, A_0 y_0\|_T \\ &\leq \|A_{n-1}\|_T \cdot \|A_{n-2}\|_T, \ldots, \|A_0\|_T \|y_0\|_T \\ &\leq \|y_0\|_T. \end{aligned}$$

In fact, for boundedness it is sufficient that there exist a T-norm such that $\|A_n\|_T \leq 1 + \gamma_n$ $\forall n$ with $\sum_0^\infty \gamma_n < \infty$. For then

$$\|y_n\|_T \leq \prod_{j=0}^{n-1} (1+\gamma_j)\|y_0\|_T \leq \exp\left(\sum_{j=0}^{n-1} \gamma_j\right)\|y_0\|_T .$$

Our previous results regarding the spectral radius no longer hold since $\rho(A)$ is not submultiplicative. The following example illustrates the problem.

Example 4.1. Let

$$A_{2j} = \begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix}, \quad A_{2j+1} = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}, \quad A_{2j+1}A_{2j} = \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}, \quad j = 0,1,\ldots .$$

We have $\rho(A_{2j+1}A_{2j}) = 4$ even though $\rho(A_{2j}) = \rho(A_{2j+1}) = 0$. The difference equation $y_{n+1} = A_n y_n$ has unbounded solutions in spite of the fact that $\rho(A_n) = 0$ $\forall n$. Note that according to Theorem 3.2, for any $\varepsilon > 0$ there are matrices $T_0(\varepsilon)$ and $T_1(\varepsilon)$ such that $\|A_{2j}\|_{T_0(\varepsilon)} < \varepsilon$, and $\|A_{2j+1}\|_{T_1(\varepsilon)} < \varepsilon$. However, in this example we cannot find a single $T(\varepsilon)$ such that $\|A_{2j}\|_{T(\varepsilon)} < \varepsilon$ and $\|A_{2j+1}\|_{T(\varepsilon)} < \varepsilon$ simultaneously unless $\varepsilon \geq 2$, since

$$4 = \rho(A_{2j+1}A_{2j}) \leq \|A_{2j+1}A_{2j}\|_{T(\varepsilon)} \leq \|A_{2j+1}\|_{T(\varepsilon)} \|A_{2j}\|_{T(\varepsilon)} < \varepsilon^2 .$$

■

This example also shows that in general it is not enough to have $\|A_n\| < 1 \, \forall n$ if the norm used to measure A_n also varies with n. It turns out that we can allow the definition of the norm to vary, but in order to guarantee boundedness the norm must vary sufficiently slowly. The following theorem gives a typical result.

Theorem 4.1. Suppose that for the difference equation (4.5) we can find a sequence of nonsingular matrices $\{T_n\}$ such that

(4.6a) $\quad \|A_n\|_{T_n} \le 1 + \gamma_n$ for $n = 0,1,\dots$ with $\sum_0^\infty \gamma_\nu < \infty$,

(4.6b) $\quad \|T_n\| < C$ where C is a constant independent of n,

(4.6c) $\quad \|T_n^{-1} T_{n-1}\| \le 1 + \beta_n$ for $n = 1,2,\dots$ with $\sum_0^\infty \beta_\nu < \infty$, $(\beta_0 := 0)$

Then y_n is bounded as $n \to \infty$.

Proof. We have that

$$\begin{aligned}
\|y_{n+1}\|_{T_n} &= \|A_n y_n\|_{T_n} \\
&\le \|A_n\|_{T_n} \|y_n\|_{T_n} \\
&\le (1+\gamma_n) \|y_n\|_{T_n} \\
&\le (1+\gamma_n) \|T_n^{-1} T_{n-1}\| \cdot \|y_n\|_{T_{n-1}} \\
&\le (1+\gamma_n)(1+\beta_n) \|y_n\|_{T_{n-1}} \\
&\le e^{\gamma_n + \beta_n} \|y_n\|_{T_{n-1}} .
\end{aligned}$$

So, by induction

$$\|y_{n+1}\|_{T_n} \le \exp \sum_{\nu=0}^{n} (\gamma_\nu + \beta_\nu) \cdot \|y_0\|_{T_0} \le K$$

Hence, for all n,

$$\|y_{n+1}\| \le \|T_n\| \cdot \|y_{n+1}\|_{T_n} \le C \cdot K,$$

which proves the theorem. ∎

We will mention some variations of this theorem. The condition (4.6b) can be replaced by

(4.6b') $\quad \|T_{n-1}^{-1} T_n\| \le 1 + \alpha_n$ with $\sum_0^\infty \alpha_\nu < \infty$

For then we have

$$\|T_n\| = \|T_{n-1}T_{n-1}^{-1}T_n\|$$
$$\leq \|T_{n-1}\| \cdot \|T_{n-1}^{-1}T_n\| \leq \|T_{n-1}\| \exp(\alpha_n)$$

and hence, by induction, we have for all n,

$$\|T_n\| \leq \|T_0\| \cdot \exp\left(\sum_0^{\infty} \alpha_\nu\right).$$

Suppose now that we use working with the ℓ_2-norm, $\|y\|^2 \equiv y^H y$. If we let $G_n := T_n^{-H} T_n^{-1}$ so that $\|y\|_{T_n} = y^H G_n y$, then the conditions (4.6b') and (4.6c) are equivalent to the single condition

$$(4.7) \qquad (1+\alpha_n)^{-1} G_n \leq G_{n+1} \leq (1+\beta_n) G_n \text{ with } \sum_0^{\infty} \alpha_\nu < \infty,\ \sum_0^{\infty} \beta_\nu < \infty.$$

where $A \leq B$ is interpreted to mean $x^H A x \leq x^H B x\ \forall x$. We leave the proof of this as an exercise.

We now turn to systems of differential equations with variable coefficients

$$(4.8) \qquad y'(t) = A(t)y(t).$$

The situation here is analogous to the difference equation case. By the inequality (4.3), $\int_0^{\infty} \mu(A(\tau))d\tau < \infty$ is sufficient for boundedness, and a fortiori $\mu(A(t)) < 0\ \forall t$ is a simple sufficient condition. It also follows that $y \to 0$ as $t \to \infty$ if $\int_0^{\infty} \mu(A(\tau))d\tau = -\infty$.

The spectral condition $\alpha(A(t)) < 0\ \forall t \geq 0$ is not sufficient. We are forced with the same problem here as in example 4.1, namely that the transformation T which relates $\alpha(A(t))$ to $\mu(A(t))$ according to Theorem 3.2 is now dependent on t. For simplicity, assume for the moment that $A(t)$ is diagonalizable for all t, $T^{-1}(t)A(t)T(t) = D(t)$ and that $T(t)$ is differentiable. Set $z(t) := T^{-1}(t)y(t)$. Then

$$(4.10) \qquad \begin{aligned} z' &= T^{-1}y' + (T^{-1})'y \\ &= (T^{-1}AT + (T^{-1})'T)z \\ &= (D - T^{-1}T')z. \end{aligned}$$

The contribution from the term $T^{-1}T'$ can be so large that $\|z(t)\| \to \infty$ as $t \to \infty$ even if the elements of $D(t)$ have negative real parts.

As in the case of difference equations, we can obtain results using a norm which varies with time, but again it must vary sufficiently slowly. The following theorem is completely analogous to Theorem 4.1.

Theorem 4.2. Suppose that for the differential equation (4.8) we can find a time-dependent differentiable transformation $T(t)$ such that

$$\text{(4.11a)} \qquad \mu_{T(t)}(A(t)) \le \gamma(t) \text{ with } \int_0^\infty \gamma(t)dt < \infty,$$

$$\text{(4.11b)} \qquad \|T(t)\| \le C \text{ where } C \text{ is a constant independent of } t,$$

$$\text{(4.11c)} \qquad \mu(-T^{-1}(t)T'(t)) \le \beta(t) \text{ with } \int_0^\infty \beta(t)dt < \infty.$$

Then $y(t)$ is bounded as $t \to \infty$.

Proof. Let $z(t) := T^{-1}(t)y(t)$ and $D(t) := T^{-1}(t)A(t)T(t)$. Then $\|y(t)\|_{T(t)} = \|z(t)\|$ and $z(t)$ satisfies the differential equation (4.10). Furthermore, $\mu_{T(t)}(A(t))$. According to (4.4), $\|z(t)\|$ can be bounded by

$$\|z(t)\| \le \|z(0)\| \exp\left[\int_0^t \mu(D(\tau) - T^{-1}(\tau)T'(\tau))d\tau\right].$$

By the subadditivity of μ,

$$\mu(D(\tau) - T^{-1}(\tau)T'(\tau)) \le \mu(D(\tau)) + \mu(-T^{-1}(\tau)T'(\tau)) \le \gamma(t) + \beta(t)$$

and so, by the conditions imposed on γ and β,

$$\|z(t)\| \le \|z(0)\| \exp\left[\int_0^t (\gamma(\tau) + \beta(\tau))d\tau\right] \le K \ \forall t > 0$$

where K is some constant independent of t. Finally, then,

$$\|y(t)\| \le \|T(t)\| \, \|z(t)\| \le CK \quad \forall t > 0$$

which completes the proof.

∎

Once again we can rewrite the conditions of the theorem in several ways. First, by (3.2c) condition (4.11c) can be replaced by a simpler and more stringent condition

$$\|T^{-1}T'\| \le \beta(t) \text{ with } \int_0^\infty \beta(t)dt < \infty$$

The condition (4.11b) can be replaced by

$$\text{(4.11b')} \qquad \mu(T^{-1}(t)T'(t)) \le \alpha(t) \text{ with } \int_0^\infty \alpha(t)dt < \infty$$

since this implies the boundedness of $\|T(t)\|$, as we now show. Note the relation of this proof to the proof of the sufficiency of condition (4.6b') in Theorem 4.1. We have

$$
\begin{aligned}
\|T(t+h)\| &= \|T(t)^{-1}(t)T(t+h)\| \\
&\le \|T(t)\|\cdot\|T^{-1}(t)[T(t)+hT'(t)+O(h^2)]\| \\
&\le \|T(t)\|\cdot\|I+hT^{-1}(t)T'(t)\| + O(h^2).
\end{aligned}
$$

We thus obtain (for $h>0$),

$$\frac{\|T(t+h)\|-\|T(t)\|}{h} \le \|T(t)\|\cdot\left[\frac{\|I+hT^{-1}(t)T'(t)\|-1}{h}\right] + O(h),$$

and in the limit, as $h\downarrow 0$,

$$\frac{d}{dt}\|T(t)\| \le \|T(t)\|\cdot\mu(T^{-1}(t)T'(t)).$$

So, by (4.4),

$$\|T(t)\| \le \|T(0)\|\exp\left(\int_0^\infty \alpha(\tau)d\tau\right).$$

If our underlying norm is the ℓ_2-norm, $\|y\|^2 = y^Hy$, and we let $G(t) := T^{-H}(t)T^{-1}(t)$, then (4.11a) is equivalent to

(4.11a') $\qquad G(t)A(t) + A^H(t)G(t) \le \gamma(t)I \;\forall t$ with $\int_0^\infty \gamma(t)dt < \infty$.

Furthermore, (4.11b') and (4.11c) can be replaced by the single equivalent condition

(4.12) $\qquad -\alpha(t)G(t) \le G'(t) \le \beta(t)G(t)$ with $\int_0^\infty \alpha(t)dt < \infty$, $\int_0^\infty \beta(t)dt < \infty$.

We will show that (4.11c) is equivalent to $G' \le \beta G$. The other half is proved similarly. Dropping the dependence on t from the notation, we see that

$$G' = (T^{-H})'T^{-1} + T^{-H}(T^{-1})' = -T^{-H}[(T^H)'T^{-H} + T^{-1}T']T^{-1}$$

and so $G' \le \beta G$ is equivalent to

$$-T^{-H}[(T^H)'T^{-H} + T^{-1}T']T^{-1} \le \beta T^{-H}T^{-1}$$

or

$$-(T^{-1}T')^H - T^{-1}T' \le \beta I,$$

which is the same as

$$\mu(-T^{-1}T') \le \beta.$$

Exercises.

4.1) Show that every square matrix A has a *Schur decomposition* of the form

$$A = U^H R U$$

where R is upper triangular and U is a unitary matrix,

$$U^H U = I$$

by starting with the Jordan canonical form $A = TJT^{-1}$ and orthogonalizing the columns of T by the Gram-Schmidt process.

4.2) The matrix A is called *normal* if $A^H A = AA^H$. Note in particular that all Hermitian matrices are normal. Prove the following:

a) if A is normal then A is diagonalizable and hence has no defective eigenvalues.
Hint: use the Schur decomposition of problem 1,

b) A is normal iff it has a complete set of mutually orthogonal eigenvec

c) if A is normal then $\rho(A) = \|A\|_2$ and $\alpha(A) = \mu_2(A)$,

d) conclude that if A is a normal matrix, $\rho(A) < 1$ is a necessary *and sufficient* condition for the boundedness of the solution of $y_{n+1} = Ay_r$
Similarly, $\alpha(A) \leq 0$ is a necessary *and sufficient* condition for the boundedness of the solution of $y' = Ay$.

4.3) Spectral conditions can be used for variable coefficient problems if all matrices involved are normal. Show that:

a) if A_n is a normal matrix for $n = 0,1,2,\ldots$ and $\rho(A_n) < 1 + \gamma_n$ with $\sum_0^\infty \gamma_n < \infty$, then all solutions of $y_{n+1} = A_n y_n$ are bounded,

b) if $A(t)$ is a normal matrix for all $t \geq 0$ and $\int_0^\infty \alpha(A(\tau))d\tau < \infty$ then all solutions of $y'(t) = A(t)y(t)$ are bounded for $t \geq 0$.

4.4) Show that the conditions (4.6b') and (4.6c) are equivalent to the single condition (4.7) when inner product norms are used.
Hint: $\|B\|_2 \leq \beta \iff B^H B \leq \beta I$.

5. Limit operations for matrices and matrix-valued functions

A sequence of p×q matrices $A_1, A_2, \ldots$, where $A_n = \left\{a_{ij}^{(n)}\right\}$, is said to converge to A^*, $\lim_{n\to\infty} A_n = A^* = [a^*_{ij}]$ if $a_{ij}^{(n)} \to a^*_{ij}$ for $1 \le i \le p$, $1 \le j \le q$.

An infinite sum of matrices is defined in the following way:

$$\sum_{n=0}^{\infty} B_n = \lim_{N\to\infty} \sum_{n=0}^{N} B_n .$$

By the triangle inequality it follows that if $\sum_{n=0}^{\infty} \|B_n\|$ is convergent, then so is $\sum_0^\infty B_n$ (although the converse is not necessarily true). In the same manner we can define $\lim_{z\to\infty} A(z)$, $A'(z)$, etc. for *matrix-valued functions of a real or complex variable*.

A power series $\sum_0^\infty B_n z^n$; $z \in \mathbb{C}$, has a *circle of convergence* in the z-plane which is equal to the smallest of the p·q circles of convergence corresponding to the series for the matrix elements. In the interior of the convergence circle, formal operations such as termwise differentiation and integration (with respect to z) are valid for the element series and therefore also for the matrix series.

Example 5.1. The matrix e^{At} can be defined by the series expansion

$$e^{At} = I + At + \frac{A^2t^2}{2!} + \frac{A^3t^3}{3!} + \ldots \tag{5.1}$$

This converges for all A and t, by the convergence of $\|I\| + \|At\| + \frac{\|At\|^2}{2!} + \ldots$ The series can thus be termwise differentiated everywhere:

$$\frac{d}{dt}(e^{At}) = A + A^2t + A^3\frac{t^6}{2!} + \ldots = Ae^{At}.$$

By applying the uniqueness theorem for the solution of the initial value problem, we obtain the following theorem.

Theorem 5.1. The solution of the initial-value problem

$$\frac{dY}{dt} = AY,\ Y(0) = I,$$

where A is a constant matrix, is $Y(t) = e^{At}$, where e^{At} is defined by (5.1).

Corollary. The solution of the initial-value problem

$$\frac{dy}{dt} = Ay,\ y(0) = c,$$

where A is a constant matrix, is $y(t) = e^{At}c$.

The representation of e^A raises the question of how to define *analytic functions of matrices* in general. (This is not the same as a matrix-valued function of a complex variable.) We first define an *analytic function of a Jordan block*, $J_i = \lambda_i I + S$, of order q, by the Taylor series

$$(5.2) \qquad f(J_i) = \sum_{p=0}^{q-1} \frac{f^{(p)}(\lambda_i)S^p}{p!}$$

Definition. Suppose that f(z) is regular for $z \in D \subset \mathbb{C}$, where D is a simply connected region which contains the spectrum of A in its interior. Let $A = T \cdot \text{blockdiag}(J_i) \cdot T^{-1}$ be the Jordan Canonical form of A. Then (5.2) makes sense for each block J_i and we define

$$f(A) := T \cdot \text{blockdiag}(f(J_i)) \cdot T^{-1}.$$

With this definition, the theory of analytic functions of *one* matrix variable closely follows the theory of analytic functions of *one* complex variable.

If $\lim_{n\to\infty} f_n(z) = f(z)$ for $z \in D$, then $\lim_{n\to\infty} f_n(J_i) = f(J_i)$ and hence $\lim_{n\to\infty} f_n(A) = f(A)$, if $\text{spectrum}(A) \subset \text{int}(D)$. This allows us to deal with operations involving limit processes. The following important theorem can then be obtained.

Theorem 5.2. All identities which hold for functions of *one* complex variabl z for $z \in D \subset C$, also hold for functions of *one* matrix variable A, if the spectrum of A is contained in int(D). The identities also hold if A has eigenvalues on ∂D, provided that these are not defective. Here D is again a simply-connected region. ∎

Thus, for example,

$$\cos^2 A + \sin^2 A = I, \ \forall A$$

$$\log(I - A) = -\sum_{n=1}^{\infty} \frac{A^n}{n} \quad \text{if } \rho(A) < 1$$

$$\int_0^\infty e^{-st} e^{At} dt = (sI - A)^{-1} \quad \text{if} \quad \alpha(A) < \text{Re } s.$$

It follows from the theorem that the series definition for e^{At} in example 5.1 is in keeping with the general definition of matrix functions. Furthermore, for example,

$$(5.3) \qquad e^{A(t+s)} = e^{At} \cdot e^{As}, \ (e^{At})^{-1} = e^{-At}; \quad t,s \in \mathbb{C}.$$

Observe also that, for two arbitrary analytic functions f and g which satisfy the conditions of the definition,

(5.4) $\qquad f(A)\cdot g(A) = g(A)\cdot f(A)$

Warning: When several non-commutative matrices are involved, one can no longer use the usual formulas for analytic functions.

Example. $(A+B)^2 - (A^2+2AB+B^2) = BA - AB$

Theorem 5.3. $e^{(A+B)t} = e^{At}\cdot e^{Bt}$ for all t, if and only if $BA = AB$.

Proof. We have

$$e^{(A+B)t} = \sum_{n=0}^{\infty} \frac{(A+B)^n t^n}{n!}$$

$$e^{At}\cdot e^{Bt} = \sum_{p=0}^{\infty} \frac{A^p t^p}{p!}\cdot \sum_{q=0}^{\infty} \frac{B^q t^q}{q!} = \sum_{n=0}^{\infty} \frac{t^n}{n!} \sum_{p=0}^{n} \frac{n!}{p!(n-p)!} A^p B^{n-p}.$$

The difference between the coefficients of $t^2/2!$ in the two expressions is $(A+B)^2 - (A^2+2AB+B^2) = BA - AB \neq 0$ if $BA \neq AB$. Conversely, if $BA = AB$, then it follows by induction that the binomial theorem holds for $(A+B)^n$, i.e.

$$(A+B)^n = \sum_{p=0}^{n} \frac{n!}{p!(n-p)!} A^p B^{n-p}$$

and hence $e^{(A+B)t} = e^{At}\cdot e^{Bt}$.

∎

Theorem 5.4.

a) $e^{At} \to 0$ as $t\to\infty$ iff all eigenvalues of A have negative real part.

b) e^{At} is bounded as $t\to\infty$ iff no eigenvalue of A has positive real part and there are no defective eigenvalues on the imaginary axis.

The proof is analogous with the proof of Theorem 1.1. Note that $e^{J_i t} = e^{\lambda_i t} e^{St}$. Alternatively, the theorem follows from the corollary to Theorem 5.1 and the results of Section 4. ∎

For linear systems with *variable coefficients*,

$$\frac{dy}{dt} = A(t)y, \quad y(0) = c,$$

the solution can no longer be written in terms of matrix exponentials, except in the commutative case in which $A(t)A'(t) = A'(t)A(t)$ for all t. It is a rather lengthy exercise to show that

$$y(t+h) - \exp\left[\int_t^{t+h} A(\tau)d\tau\right]y(t) \sim \frac{h^3}{12}[A'(t)A(t) - A(t)A'(t)]y(t) \quad (\text{as } h\to 0)$$

6. Stable families of matrices

Let us now return to difference equations with constant coefficients,

(6.1) $$y_{n+1} = Ay_n.$$

In applications where (6.1) arises as a discretization of some differential equation, the iteration matrix A will depend on a stepsize h. In this case it is natural to consider the following question: when does the *family* of problems $\{y_{n+1} = A(h)y_n\}$ have solutions which are bounded for all n independently of h? The most important applications of this concept are to partial differential equations. See Chapter 4 of Richtmyer and Morton [4] for more on this. This section is based on their treatment of the material.

Definition. A family F of matrices is said to be *stable* if there exists a constant C such that for all $A \in F$ and all $\nu = 1,2,3,\dots$.

(6.2) $$\|A^\nu\| \le C.$$

Note that this is stronger than simply requiring that A^ν be bounded for all A in F, since the bound C must be independent of A. It is a *necessary* condition, however, that each matrix $A \in F$ be bounded in the sense of Section 1.

The following theorem of Kreiss gives some equivalent conditions for the stability of the family F.

The Kreiss Matrix Theorem

The stability of the family F is equivalent to each of the following:

a) There exists a constant C_1 such that for all $A \in F$ and all $z \in \mathbb{C}$ with $|z| > 1$, $(A - zI)^{-1}$ exists and

(6.3) $$\|(A - zI)^{-1}\| \le \frac{C_1}{|z|-1}.$$

b) There exist constants C_2 and C_3 and, to each $A \in F$ a nonsingular matrix S such that

i) $\|S\|,\ \|S^{-1}\| \le C_2$

ii) $B := S^{-1}AS$ is upper triangular with $|b_{ij}| \le C_3 \min(1 - |b_{ii}|, 1 - |b_{jj}|)$ for $i \neq j$.

c) There exists a constant C_4 and, for each $A \in F$ a positive definite matrix G such that

$$C_4^{-1} I \le G \le C_4 I$$
$$A^H G A \le G.$$

■

The condition c) of this theorem can be rewritten in a more familiar form as

c') There exists a constant C_5 and, for each $A \in F$ a nonsingular matrix T such that

$$\|A\|_T \leq 1$$
$$\mathrm{cond}(T) \leq C_5$$

where $\|\cdot\|$ is the ℓ_2-norm and the T-norm $\|\cdot\|_T$ is as in Section 4. It is left as an exercise to show that c) and c') are equivalent. (Hint: take $G = T^{-H}T^{-1}$).

For the complete proof of the Kreiss Matrix Theorem, the reader is referred to [4].

The theorem is proved by showing that stability $\Rightarrow$ a) $\Rightarrow$ b) $\Rightarrow$ c) $\Rightarrow$ stability. Note that if the alternative formulation c') is used then the step c') $\Rightarrow$ stability is completely analogous to the proofs of Section 4. Furthermore, the proof that b) $\Rightarrow$ c') is quite similar to the proof of Theorem 3.2, except that the triangular decomposition of b) is used as the starting point in place of the Jordan Form. It can be shown that c') follows from b) by taking $T := SD$, where D is a suitably chosen diagonal scaling matrix which is fixed independent of A.

The condition a) is called the *resolvent condition*, since it gives a bound on how quickly the resolvent $(A - zI)^{-1}$ can grow as z approaches the unit circle. Note that in particular it implies that all the eigenvalues of A must lie in the closed unit disk and that any eigenvalue on the unit circle must be simple. The proof that stability $\Rightarrow$ a) is quite simple, so we will present it here. If the family F is stable, then each $A \in F$ must have eigenvalues lying in the closed unit discs, by Theorem 1.1. Hence for $|z| > 1$, $(A - zI)$ is invertible and

$$\|(A - zI)^{-1}\| = \|\sum_{\nu=0}^{\infty} A^{\nu} z^{-\nu-1}\| \leq C \sum_{\nu=0}^{\infty} |z|^{-\nu-1} = \frac{C}{|z|-1}.$$

Thus (6.3) holds with $C_1 := C$.

REFERENCES

[1] Strang, G., Linear Algebra and its Applications, Acad.Press 1976

[2] Dahlquist, G. and Björck, Å., Numerical Methods, Prentice-Hall, Englewood Cliffs 1974.

[3] Bellman, R., Introduction to Matrix Analysis, Mc Graw Hill 1960.

[4] Richtmyer, R.D. and Morton K.W., Difference Methods for Initial Value Problems, Wiley 1967.