Elemente der Mathematik

Matrices with a strictly dominant eigenvalue

Helmut Länger

Helmut Länger, born in 1951, obtained his doctoral degree in mathematics at the
Vienna University of Technology as a student of Professor Nöbauer. After spending
a year in Darmstadt he returned to Vienna in 1977 where he now holds the position
of an Associate Professor at the Institute of Algebra and Computational Mathematics
of the University of Technology. His main research interests are generalizations of
semigroups, discrete problems in biology, inequalities, algebraic structures occurring
in axiomatic quantum mechanics and universal algebra. His hobbies include hiking
and music.

1 Introduction

Matrices with a strictly dominant eigenvalue often appear in applications, e.g. when dealing with systems of linear difference or differential equations with positive coefficients. In order to study the asymptotic behaviour of such systems (when time goes to infinity) it is useful to have simple formulas at hand for $\lim_{k \to \infty} (A^k / \lambda^k)$ (resp. $\lim_{t \to \infty} (e^{At} / e^{\lambda t})$), where A denotes the corresponding matrix of coefficients and λ the strictly dominant eigenvalue of A. The aim of this paper is to provide such formulas and to show their usefulness by applying them to finite Markov chains as well as to a model in mathematical ecology.

2 Strictly dominant eigenvalues

Definition 2.1 Let n be a positive integer, A a complex $n \times n$-matrix and λ an eigenvalue of A. λ is called *discretely strictly dominant* if λ is simple and $|\lambda| > |\mu|$ for every eigenvalue μ of A different from λ. λ is called *continuously strictly dominant* if λ is simple and $\text{Re}(\lambda) > \text{Re}(\mu)$ for every eigenvalue μ of A different from λ. An n-dimensional column vector $\vec{b} \neq \vec{0}$ is called a *right eigenvector* of A corresponding

Grundlage des nachfolgenden Beitrags ist die aus der Linearen Algebra bekannte Jordansche Normalform einer quadratischen Matrix, welche sich mit Hilfe der Eigenwerte und Eigenvektoren der gegebenen Matrix bestimmen lässt. Damit findet der Autor eine elegante Berechnung des Grenzwertes der Folge der k-ten Potenzen A^k für gewisse quadratische Matrizen A. Überraschenderweise lassen sich diese Berechnungen nun auf die Übergangsmatrizen gewisser endlicher Markovscher Ketten anwenden. Dies ermöglicht eine einfache Bestimmung des Endzustandes eines solchen Prozesses. *jk*

to λ if $A\vec{b} = \lambda\vec{b}$. An n-dimensional row vector $\vec{c} \neq \vec{0}$ is called a *left eigenvector* of A corresponding to λ if $\vec{c}A = \lambda\vec{c}$.

The aim of this paper is to prove the following theorem and to sketch some of its applications (for similar results cf. e.g. [10]):

Theorem 2.1 *Let A be a complex square matrix, λ a simple eigenvalue of A, $\vec{b}$ a corresponding right eigenvector and $\vec{c}$ a corresponding left eigenvector. Then $\vec{c}\vec{b} \neq 0$ and the following hold:*

(i) *If λ is discretely strictly dominant and $\lambda \neq 0$ then $\displaystyle\lim_{k \to \infty} (A^k/\lambda^k) = \vec{b}\vec{c}/\vec{c}\vec{b}$.*

(ii) *If λ is continuously strictly dominant then $\displaystyle\lim_{t \to \infty} (e^{At}/e^{\lambda t}) = \vec{b}\vec{c}/\vec{c}\vec{b}$.*

(Here and in the following the vectors $\vec{b}$ and $\vec{c}$ are also interpreted as matrices and 1×1-matrices are identified with the corresponding scalar.)

Proof. Let $J = \mathrm{diag}(\lambda, J_2, \ldots, J_m)$ be the Jordan normal form of A where for $i = 2, \ldots, m$ J_i is the $n_i \times n_i$-matrix

$$\begin{pmatrix} \lambda_i & 1 & 0 & \ldots & 0 \\ 0 & \lambda_i & 1 & \ddots & \vdots \\ 0 & 0 & \lambda_i & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \ldots & 0 & 0 & \lambda_i \end{pmatrix} .$$

Then there exists a regular $n \times n$-matrix B with $A = BJB^{-1}$. Let $\vec{f}$ denote the first column of B and $\vec{g}$ the first row of B^{-1}. Since B and B^{-1} are regular, $\vec{f}, \vec{g} \neq \vec{0}$. Because of $AB = BJ$, $\vec{f}$ is a right eigenvector of A corresponding to λ and because of $B^{-1}A = JB^{-1}$, $\vec{g}$ is a left eigenvector of A corresponding to λ. Hence there exist $\alpha, \beta \in \mathbf{C} \setminus \{0\}$ with $\vec{f} = \alpha\vec{b}$ and $\vec{g} = \beta\vec{c}$. Since $B^{-1}B = I$, we have $\vec{g}\vec{f} = 1$. This shows $\vec{c}\vec{b} = (\vec{g}\vec{f})/(\alpha\beta) = 1/(\alpha\beta) \neq 0$ whence $\alpha\beta = 1/(\vec{c}\vec{b})$.

(i) If λ is discretely strictly dominant and $\lambda \neq 0$ then

$$\lim_{k \to \infty} \frac{J_i^k}{\lambda^k}$$

$$= \lim_{k \to \infty} \begin{pmatrix} (\lambda_i/\lambda)^k & \binom{k}{1}(1/\lambda)(\lambda_i/\lambda)^{k-1} & \ldots & \binom{k}{n_i-1}(1/\lambda^{n_i-1})(\lambda_i/\lambda)^{k-n_i+1} \\ 0 & (\lambda_i/\lambda)^k & \ddots & \vdots \\ \vdots & \ddots & \ddots & \binom{k}{1}(1/\lambda)(\lambda_i/\lambda)^{k-1} \\ 0 & \ldots & 0 & (\lambda_i/\lambda)^k \end{pmatrix}$$

$$= O$$

for $i = 2, \ldots, m$ since $|\lambda_i/\lambda| < 1$ and hence

$$\lim_{k \to \infty} \frac{A^k}{\lambda^k} = \lim_{k \to \infty} B \frac{J^k}{\lambda^k} B^{-1}$$

$$= \lim_{k \to \infty} B \operatorname{diag}(1, \frac{J_2^k}{\lambda^k}, \ldots, \frac{J_m^k}{\lambda^k}) B^{-1} = B \operatorname{diag}(1, O, \ldots, O) B^{-1}$$

$$= \vec{f}\vec{g} = \alpha\beta\vec{b}\vec{c} = \frac{\vec{b}\vec{c}}{\vec{c}\vec{b}}.$$

(ii) If λ is continuously strictly dominant then

$$\lim_{t \to \infty} \frac{e^{J_i t}}{e^{\lambda t}}$$

$$= \lim_{t \to \infty} \begin{pmatrix} \exp((\lambda_i - \lambda)t) & (t/1!)\exp((\lambda_i - \lambda)t) & \ldots & (t^{n_i - 1}/(n_i - 1)!)\exp((\lambda_i - \lambda)t) \\ 0 & \exp((\lambda_i - \lambda)t) & \ddots & \vdots \\ \vdots & \ddots & \ddots & (t/1!)\exp((\lambda_i - \lambda)t) \\ 0 & \ldots & 0 & \exp((\lambda_i - \lambda)t) \end{pmatrix}$$

$$= O$$

for $i = 2, \ldots, m$ since $\operatorname{Re}(\lambda_i - \lambda) < 0$ and hence

$$\lim_{t \to \infty} \frac{e^{At}}{e^{\lambda t}} = \lim_{t \to \infty} B \frac{e^{Jt}}{e^{\lambda t}} B^{-1} = \lim_{t \to \infty} B \operatorname{diag}(1, \frac{e^{J_2 t}}{e^{\lambda t}}, \ldots, \frac{e^{J_m t}}{e^{\lambda t}}) B^{-1}$$

$$= B \operatorname{diag}(1, O, \ldots, O) B^{-1} = \vec{f}\vec{g} = \alpha\beta\vec{b}\vec{c} = \frac{\vec{b}\vec{c}}{\vec{c}\vec{b}}. \qquad \square$$

Corollary 2.1 *Under the assumptions of Theorem* 2.1 *we have*

$$\lim_{k \to \infty} \frac{A^k \vec{d}}{\lambda^k} = \frac{\vec{c}\vec{d}}{\vec{c}\vec{b}} \vec{b} \qquad \left(resp. \ \lim_{t \to \infty} \frac{e^{At} \vec{d}}{e^{\lambda t}} = \frac{\vec{c}\vec{d}}{\vec{c}\vec{b}} \vec{b} \right)$$

for every complex column vector $\vec{d}$ of the same dimension as A since $(\vec{b}\vec{c})\vec{d} = \vec{b}(\vec{c}\vec{d}) = (\vec{c}\vec{d})\vec{b}$.

Definition 2.2 A real *matrix* is called *non-negative* (resp. *positive*) if all of its elements are non-negative (resp. positive). We say that a non-negative square matrix A satisfies condition (R) if there exists a positive integer m such that A^m is positive.

We now restate the fundamental theorem on non-negative matrices (cf. [9] and [3]):

Theorem 2.2 (Frobenius-Perron Theorem) (cf. e.g. [7, p. 191]) *Every non-negative square matrix satisfying condition* (R) *has a positive discretely strictly dominant eigenvalue and a corresponding positive right eigenvector.*

For obtaining bounds for the discretely strictly dominant eigenvalue of a non-negative square matrix satisfying condition (R) the following well-known result is useful (for the sake of completeness we provide a short proof, for similar results cf. [1]):

Lemma 2.1 *Every real eigenvalue of a real square matrix, to which there exists a non-negative right eigenvector, lies between the smallest and greatest column sum of the matrix.*

Proof. Assume $A = (a_{ij})_{i,j=1,\ldots,n}$ to be a real matrix, λ to be a real eigenvalue of A and $\vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ to be a non-negative right eigenvector of A corresponding to λ. Then $\sum_{j=1}^{n} a_{ij} b_j = \lambda b_i$ for $i = 1, \ldots, n$ and hence

$$\lambda \sum_{i=1}^{n} b_i = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} b_j = \sum_{j=1}^{n} b_j \sum_{i=1}^{n} a_{ij} \in \left[\sum_{j=1}^{n} b_j \min_{k=1,\ldots,n} \sum_{i=1}^{n} a_{ik}, \sum_{j=1}^{n} b_j \max_{k=1,\ldots,n} \sum_{i=1}^{n} a_{ik} \right].$$

Division by $\sum_{i=1}^{n} b_i$ completes the proof. $\square$

To a non-negative square matrix we now assign a directed graph:

Definition 2.3 For a non-negative square matrix $A = (a_{ij})_{i,j=1,\ldots,n}$ let $G(A)$ denote the directed graph with vertex-set $\{1, \ldots, n\}$ having a directed edge from vertex i to vertex j if and only if $a_{ij} > 0$ $(i, j \in \{1, \ldots, n\})$. A directed graph is called *strongly connected* (cf. e.g. [4]) if from any of its vertices there exists a directed walk to every other one of its vertices.

For checking property (R) the following result is useful:

Theorem 2.3 (cf. [5]) *A non-negative square matrix A satisfies condition* (R) *if and only if $G(A)$ is strongly connected and in $G(A)$ there exist two closed directed walks with coprime lengths. (Two integers are said to be coprime to each other if their greatest common divisor equals* 1.*)*

Example 2.1 If $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ then $G(A)$ is strongly connected and all closed directed walks have even length. Hence there do not exist closed directed walks with coprime lengths. This is in accordance with the fact that for $k \geq 1$, A^k either equals $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ or $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. If $A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ then $G(A)$ is strongly connected since $1, 2, 1$ and $1, 2, 3, 1$ are closed directed walks and $\gcd(2, 3) = 1$. Hence Theorem 2.3 is applicable. Indeed, $A^5 = \begin{pmatrix} 2 & 1 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

Now we will mention two applications of our result:

3 Regular Markov chains

Consider a system which at time $0, 1, 2, \ldots$ is in exactly one of the states $1, \ldots, n$. For $i, j = 1, \ldots, n$ and $k \geq 0$ let b_{jk} denote the probability that the system is in state j at time k and a_{ij} denote the probability that the system is in state i provided that it was in state j one time-unit before. Hence the state of the system only depends on the state one time-unit before (and not on the state at earlier times) and the transition probabilities are constant. This process is a finite Markov chain with transition matrix $A = (a_{ij})_{i,j=1,\ldots,n}$ and state vector $\vec{b}_k = \begin{pmatrix} b_{1k} \\ \vdots \\ b_{nk} \end{pmatrix}$ at time k. We have $a_{ij} \geq 0$, $\sum_{i=1}^{n} a_{ij} = 1$ (such a matrix is usually called stochastic), $b_{jk} \geq 0$ and $\sum_{j=1}^{n} b_{jk} = 1$. According to the Theorem of Total Probability (cf. e.g. [2, p. 56]) one obtains $\vec{b}_{k+1} = A\vec{b}_k$ for all $k \geq 0$. The Markov chain is called *regular* if A satisfies condition (R). Now we have the following well-known theorem (for another proof of this theorem cf. e.g. [6]):

Theorem 3.1 *The state vectors of a regular Markov chain converge to the unique right eigenvector of the corresponding transition matrix with component sum* 1 *corresponding to the eigenvalue* 1.

Proof. Assume A to be the transition matrix corresponding to a regular Markov chain. Because of Theorem 2.2 A has a positive discretely strictly dominant eigenvalue and a corresponding positive right eigenvector. Since all the column sums of A are 1, this eigenvalue is 1 according to Lemma 2.1. Because of Theorem 2.1 A^k and hence also $\vec{b}_k$ converges. Since multiplication of matrices is continuous, $\vec{b}_k$ converges to the unique right eigenvector of A with component sum 1 corresponding to the eigenvalue 1.   □

4 The Leslie model

This model is a fundamental model in mathematical ecology. It is used when considering populations which are divided into several age classes (cf. e.g. [8]).

Consider a fixed population divided into age classes $1, \ldots, n$ each of which is of the length of a certain time unit. For $j = 1, \ldots, n$ let a_{1j} denote the average number of offspring of an individual chosen at random in age class j and for $i = 2, \ldots, n$ let $a_{i,i-1}$ denote the probability that an individual chosen at random in age class $i - 1$ reaches age class i. Put $a_{ij} := 0$ otherwise. For $i = 1, \ldots, n$ and $k \geq 0$ let d_{ik} denote the size of age class i at time k. $\vec{d}_k := \begin{pmatrix} d_{1k} \\ \vdots \\ d_{nk} \end{pmatrix}$ is called the age class distribution at time k. It follows $\vec{d}_{k+1} = A\vec{d}_k$ for $k \geq 0$.

Theorem 4.1 *If there exist coprime $i, j \in \{1, \ldots, n\}$ with $a_{1i}, a_{1j} > 0$ (this means that the i-th and j-th age class have positive birth rate), if $a_{21}, a_{32}, \ldots, a_{n,n-1} > 0$ (this means that all survival rates from one age class to the next are positive) and if there*

exist $s, t \in \{1, \ldots, n\}$ *with* $s \leq t$ *and* $d_{s0}, a_{1t} > 0$ *then A has a positive discretely strictly dominant eigenvalue λ_0 and* (i) *and* (ii) *hold:*

(i) $\lim\limits_{k \to \infty} (d_{ik}/d_{jk}) = b_i/b_j$ *for all* $i, j = 1, \ldots, n$.

(ii) $\lim\limits_{k \to \infty} (d_{i,k+1}/d_{ik}) = \lambda_0$ *for all* $i = 1, \ldots, n$.

Here $\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ *denotes a positive right eigenvector of A corresponding to λ_0.*

Remark 4.1 The following proof shows that almost all $\vec{d}_k$'s are positive. Because of (i), $\vec{b} / \sum\limits_{i=1}^{n} b_i$ is called the *stable age distribution* and because of (ii), λ_0 is called the *natural growth rate* of the considered population, respectively.

Proof of Theorem 4.1. Let m denote the greatest index with $a_{1m} > 0$, such that the matrix A has the structure

$$
A = \begin{pmatrix}
a_{11} & \cdots & \cdots & a_{1m} & 0 & \cdots & 0 \\
a_{21} & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & \ddots & \ddots & & & & \vdots \\
\vdots & \ddots & \ddots & \ddots & & & \vdots \\
\vdots & & \ddots & \ddots & \ddots & & \vdots \\
\vdots & & & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & \cdots & \cdots & 0 & a_{n,n-1} & 0
\end{pmatrix}.
$$

We then let B be the upper left $m \times m$ block of A and $\vec{f}_k := \begin{pmatrix} d_{1k} \\ \vdots \\ d_{mk} \end{pmatrix}$ for $k \geq 0$. Since for every $s, t \in \{1, \ldots, m\}$, $s, s-1, \ldots, 1, m, m-1, \ldots, t$ is a directed walk from s to t and $1, i, i-1, \ldots, 1$ and $1, j, j-1, \ldots, 1$ are two closed directed walks with coprime lengths i and j, respectively, it follows from Theorems 2.3 and 2.2 that B has a positive discretely strictly dominant eigenvalue λ_0. Let $\vec{b} := \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ and $\vec{c}$ denote a positive right, respectively left eigenvector of B corresponding to λ_0 and put $\alpha := (\vec{c}\vec{f}_0)/(\vec{c}\vec{b})$. Then

$$
\lim_{k \to \infty} \frac{\vec{f}_k}{\lambda_0^k} = \lim_{k \to \infty} \frac{B^k \vec{f}_0}{\lambda_0^k} = \alpha \vec{b}
$$

according to Corollary 2.1. Now $|A - \lambda I| = (-\lambda)^{n-m} |B - \lambda I|$ which can be seen by expanding the first determinant $n - m$ times by the last column. Hence λ_0 is also a discretely strictly dominant eigenvalue of A. Now, due to the special form of A the

vectors $\vec{f}_{k+1}$ and $\vec{b}$ easily extend by

$$d_{m+1,k+1} = a_{m+1,m}d_{mk}$$

$$\vdots$$

$$d_{n,k+1} = a_{n,n-1}d_{n-1,k}$$

for $k \geq 0$ and $b_i := a_{i,i-1} \cdot \ldots \cdot a_{m+1,m}b_m/\lambda_0^{i-m}$ for $i = m+1, \ldots, n$ to the corresponding age class distribution and positive right eigenvector corresponding to λ_0 of the entire matrix A, respectively, and $\lim_{k \to \infty}(d_{ik}/\lambda_0^k) = \alpha b_i$ for $i = 1, \ldots, n$. Since $\vec{b}$ is positive, almost all $\vec{d}_k$'s are positive. Now (i) and (ii) easily follow. $\square$

References

[1]   Collatz, L.: Einschließungssatz für die charakteristischen Zahlen von Matrizen. *Math. Zeitschrift* 48 (1942/1943), 221–226.

[2]   DeGroot, M.H.: *Probability and Statistics*. Addison-Wesley, Reading, Mass. 1975.

[3]   Frobenius, G.: Über Matrizen aus nicht negativen Elementen. Sitzungsber. Preuß. Akad. Wiss., Berlin 1912, 456–477.

[4]   Gross, J., and Yellen, J.: *Graph Theory and its Applications*. CRC Press, Boca Raton 1999.

[5]   Karigl, G.: On a graph theoretic description of regular Markov chains and positive linear systems. *Demonstr. Math.* 30 (1997), 937–944.

[6]   Kemeny, J.G., and Snell, J.L.: *Finite Markov Chains*. Springer, New York 1976.

[7]   Luenberger, D.G.: *Introduction to Dynamic Systems*. Wiley, New York 1979.

[8]   Nöbauer, W. und Timischl, W.: *Mathematische Modelle in der Biologie*. Vieweg, Braunschweig 1979.

[9]   Perron, O.: Zur Theorie der Matrices. *Math. Ann.* 64 (1907), 248–263.

[10] Varga, R.S.: *Matrix Iterative Analysis*. Prentice-Hall, Englewood Cliffs, New Jersey 1962.

Helmut Länger
Technische Universität Wien
Institut für Algebra und Computermathematik
Wiedner Hauptstraße 8–10
A-1040 Wien, Austria
e-mail: h.laenger@tuwien.ac.at